IMK
Institut für Makroökonomie
und Konjunkturforschung
Macroeconomic Policy Institute

May 2012 94

Working Paper

Sabine Stephan | Katja Rietzler

Monthly recession predictions in real time: A density forecast approach for German industrial production

Abstract
In this paper we present a methodology which can help to improve the assessment of the current economic situation. We propose an approach which combines multivariate single equations to forecast the monthly growth rate of industrial production with a density forecast. This allows to estimate the current recession probability. In the analysis the focus is on the real-time problem, i.e. the fact that the reference series (industrial production) as well as important indicators are not available on a timely basis and are often revised substantially over an extended period. For this reason the whole analysis is carried out under real-time conditions. Indeed the forecast of the recession probabilities allows to identify the recession well before it can be seen in the official data. This result is encouraging. But there is still a substantial need for further research.

Keywords: recession probability, density forecast, forecasting, business cycleresearch, real-time data, real-time conditions

JEL Codes: E32, E37, C53

Monthly recession predictions in real time: A density forecast approach for German industrial production*

Sabine Stephan**
Macroeconomic Policy Institute
(IMK), Duesseldorf
Sabine-Stephan@boeckler.de

Katja Rietzler
Rietzler Economics, Berlin
info@rietzler-economics.eu

May 22, 2012

Abstract

In this paper we present a methodology which can help to improve the assessment of the current economic situation. We propose an approach which combines multivariate single equations to forecast the monthly growth rate of industrial production with a density forecast. This allows to estimate the current recession probability. In the analysis the focus is on the real-time problem, i.e. the fact that the reference series (industrial production) as well as important indicators are not available on a timely basis and are often revised substantially over an extended period. For this reason the whole analysis is carried out under real-time conditions. Indeed the forecast of the recession probabilities allows to identify the recession well before it can be seen in the official data. This result is encouraging. But there is still a substantial need for further research.

Keywords: recession probability, density forecast, forecasting, business cycle research, real-time data, real-time conditions

JEL Codes: E32, E37, C53

*This paper is based on an expertise commissioned by the German Federal Ministry of Finance. We thank our co-authors of the expertise, Sven Schreiber, Christian Proaño Acosta and Thomas Theobald for helpful comments and extensive programming support. We also thank Christian Schumacher and Jörg Breitung for sharing the real-time data set.

**Corresponding author: Sabine Stephan, IMK, Hans-Boeckler-Str. 39, D-40476 Duesseldorf, Germany; phone +49/211/7778 335.

1 Introduction

An early and reliable identification of business cycle turning points is among the key challenges of macroeconomic research. Whereas the expected trend can be forecast quite accurately during stable expansions, when there is no change of direction, recessions which often involve a sudden and steep decline in economic activity are often detected with delay or only in retrospect. This makes a targeted and stabilising economic policy difficult.

Detecting cyclical turning points in the most recent observations is particularly challenging. On the one hand this is due to the fact that a phase which is characterised by declining economic activity has to have a certain minimum duration to be identified as a recession. Logically this means that the recession can only be reported at the end of this minimum period. On the other hand the most recent observations of key macroeconomic time series are subject to great uncertainty. This uncertainty results from the fact that indicators such as industrial production, new orders or price indices are published with a certain delay (publication lag) and that these indicators are subject to repeated and sometimes substantial revisions and many months may pass until the revision process is completed. For this reason the assessment of the economic situation resulting from these indicators in real time may differ substantially from the retrospective assessment based on final data (cf. e.g. Stark and Croushore, 2002).

The real-time problem and the implications associated with the dating algorithm have far-reaching consequences for a timely detection and determination of turning points. Suppose that industrial production, which exhibits a publication lag of two months, is used as a reference series for the business cycle and a dating procedure is applied which is based on the duration and intensity (amplitude) of the cumulative production decline over five months. If this dating algorithm reports in period t that the definition of a recession was fulfilled for the first time in period t-2 (the latest reporting period for which data are available in period t), this means that the turning point occurred at some point between t-7 and t-3. Thus, a timely detection and exact determination of turning points is almost impossible.

Rather than focussing on turning points the approach taken here is to predict whether the economy is currently in a recession. This issue is not trivial at all, because at time t there are no official data of industrial production for the previous and the current month. We are thus in the unusual situation that we have to forecast an observation, which from the point of view of t is already in the past. Or in other words: here a "back-cast" is a genuine forecast. The same is true for the "now-cast" as well. In the following we describe a two-step procedure to predict recession probabilities. In a first step we forecast the growth rate of industrial production at time *t-1*, *t*, and *t+1* and calculate cumulative growth rates over five months (the defined minimum duration of a recession). In a second step we derive the density functions of these cumulative growth rates and estimate the probability that the estimated growth falls below a certain threshold indicating a recession. The determination of the threshold is based on the triangular approximation approach proposed by Harding and Pagan (2002).

In contrast to factor models (cf. e.g. Schumacher and Breitung, 2008) and pooling approaches (cf. e.g. Drechsel and Scheufele, 2010, 2011), which extract relevant information from large data sets, we deliberately specify a small model. We apply a multivariate single equation approach and estimate the growth rate of industrial production using autoregressive terms and appropriate (leading) indicators. Compared to more complex models, which easily become "black boxes", this approach has the advantage that the estimation equation directly shows which indicators contribute to an explanation of the dependent variable and also that the results are transparent.

In contrast to other empirical studies which usually carry out their analysis only once and for a specific reference period, optimising their forecast models for the respective data set, the focus of this study is on obtaining the best possible forecast under real-time conditions with a limited input of time and work as well as extensive automation. This is a necessary requirement if one is interested in an application of this recession probability forecast on a regular basis. However, it largely determines the model specification, because

- the indicators used must be easily available,
- publication lags of variables used are taken into account,
- only information, which was actually available at the time t, is considered, and
- model specification and forecast evaluation are largely automated.

Against the background that publication lags differ for different time series (see Table 1) and in each case the maximum available information is to be used, we have to specify individual equations for each forecast horizon (*t-1*, *t*, *t+1*). Thus, the term "single equation approach" applies to each of the forecast horizons.

The paper is structured as follows: Section 2 presents the data. Section 3 describes the preliminary analyses which serve the objective to select the most promising variables for the explanation of the growth rate of industrial production from the large available data set. As a first step bivariate estimations of the growth rate of industrial production and each indicator are carried out. On the basis of these preliminary bivariate equations promising variables for a multivariate general-to-specific approach are selected (Section 4). As there are several potential specifications for each forecast horizon, these specifications are compared in terms of their in-sample forecast performance for two different sample sizes. At the end of this selection process we obtain one estimation equation for each forecast horizon. Section 4 describes which indicators are relevant in each case. The three selected specifications serve as a basis for the examination of the forecast performance of this approach. Section 5 illustrates how the recession probabilities are estimated using a density forecast approach. Section 6 presents the evaluation of the out-of-sample recession forecasts. Note, that the whole analysis presented in Section 3 to 6 is carried out under real-time conditions. Section 7 concludes with a summary assessment of the approach.

2 Data

A large number of indicators is available for economic analyses and forecasts. They include "hard" indicators of the real economy such as new orders, financial market indicators such as interest rates or indices of stocks or fixed-income securities as well as "soft" indicators such as survey results. The appropriate scope of data depends on the concrete objective of the analysis. For factor models (Forni et al., 1999, Schumacher and Breitung, 2008) or pooling (Drechsel and Scheufele, 2010, 2011) large data bases are preferable. However, the current analysis is concerned with the contribution of each individual indicator rather than the best forecast based on a maximum of indicators. The objective is a relatively parsimonious model. In this case it is important to include indicators from all relevant categories.

The existing literature on forecasts of industrial production shows diverse approaches to the selection of variables. Drechsel and Scheufele (2010) use about 90 indicators for their forecast of German industrial production. By contrast Carstensen et al. (2011) limit their data base for the analysis of forcasts of euro area industrial production to merely seven indicators. With a data base of 26 indicators this study ranges somewhere in the middle. All relevant time series are included in the data base. One aspect of data selection is the easy availability and clarity of the data. In addition, this objective here is not to carry out a one-off analysis, but rather to develop a forecast model which can be applied regularly without excessive efforts.

Table 1: Variables

Variable	Source	Abbreviation	Publ. lag (months)	Stochastic properties	Revisions
Dependent variable					
Industrial production excl. construction	Bundesbank	IPRO	2	I(1)	yes
Real economic indicators					
Orders received from the domestic market	Bundesbank	ORDERS_DOM	2	I(1)	yes
Orders received from abroad	Bundesbank	ORDERS_FOR	2	I(1)	yes
Job vacancies	Bundesbank	VAC	0	I(1)	yes*
Surveys					
Ifo index: business climate	ifo Institute	IFO_CLIMATE	0	I(0)	no
Ifo index: business situation	ifo Institute	IFO_SIT	0	I(0)	no
Ifo index: business expectations	ifo Institute	IFO_EXP	0	I(0)	no
ZEW indicator: economic sentiment	ZEW‡	ZEW_SENT	0	I(0)	no
ZEW indicator: economic situation	ZEW‡	ZEW_SIT	0	I(0)	no
Industrial confidence	DG ECFIN/Ecowin	INDCONF	1	I(0)	no
Production expectations	DG ECFIN/Ecowin	PRODEX	1	I(0)	no
Economic sentiment	DG ECFIN/Ecowin	ECSENT	1	I(0)	no
Composite indicators					
OECD Composite Leading Indicator	OECD	CLI	2	trend stat.	yes
Commerzbank Early Bird Indicator	Commerzbank	COM_EB	1	I(1)	yes*
Prices					
Consumer prices	Bundesbank	CPI	1	I(1)	yes
Real effective exchange rate	OECD	REER	2	I(1)	yes*
Brent spot price, USD/barrel	Reuters/Ecowin	OIL	0	I(1)	no
Interest rates and financial indicators					
CDAX	Bundesbank	CDAX	0	I(1)	no
Credit growth	Bundesbank	CREDIT	1	I(1)	yes*
Fixed income index	Bundesbank	FIX	0	I(1)	no
Fibor/Euribor	Bundesbank	EURIBOR	0	I(1)	no
Corporate Spread	Bundesbank, own calc.	C_SPREAD	0	I(1)	no
Term-Spread (1y-3m)	Bundesbank, own calc.	YC1Y	0	I(0)	no
Term-Spread (2y-3m)	Bundesbank, own calc.	YC2Y	0	I(0)	no
Term-Spread (3y-3m)	Bundesbank, own calc.	YC3Y	0	I(0)	no
Term-Spread (5y-3m)	Bundesbank, own calc.	YC5Y	0	I(0)	no
Term-Spread (10y-3m)	Bundesbank, own calc.	YC10Y	0	I(1)**	no

‡ Centre for European Economic Research

* Revision history not available.

** Is treated as I(0) in analogy with the other term spreads.

Table 1 provides an overview of the variables, their sources, their publication lags (i.e. the time passing between the end of the reporting period and the first release of the respective data) as well as the stochastic properties of the data. It also reveals whether the time series are subject to revisions. With the exception of the credit supply, which is available from September 1998, all time series begin in February 1993. To determine the stochastic properties of the time series ADF tests (Fuller, 1976, Dickey and Fuller, 1979, 1981) were carried out (see Table 10). Real-time data were taken from the real-time data bases of the Deutsche Bundesbank (industrial production, new orders, consumer price index) and the OECD (OECD composite leading indicator). Further details on how the real-time data were edited as well as an analysis of the revisions are provided in Schreiber et al. (2012).

3 Preliminary examinations for variable selection

3.1 Bivariate estimations

26 macroeconomic indicators were chosen for this analysis. They can be classified into three fields: real economic indicators, financial market indicators as well as surveys. The number of indicators is too large for a general-to-specific approach including all indicators. For this reason, an approach has to be identified which allows to distinguish promising indicators from unsuitable ones.

A widely applied approach consists in identifying the leads of individual indicators using a Granger causality test (Granger, 1969). This test examines whether an autoregressive estimation of variable x can be improved if lagged values of variable y (up to a certain pre-defined lag) are included. If this is the case, y is said to Granger cause x. At the same time the test analyses whether an inclusion of lagged variables of x improves an autoregressive estimation of y. In the latter case x would Granger cause y. If x Granger causes y, but y does not Granger cause x, then x is assumed to have leading indicator properties with respect to y.

The Granger causality test is not applicable for the current analysis, because it takes into account all lags starting at lag 1. Under real-time conditions, however, the first few lags of a number of variables are not available at the time of the forecast due to the publication lag. As Table 2 shows this problem is worsening with an increasing forecast horizon.

Table 2: Lag structure depending on the publication lag

Publication lags of the indicator	Estimation/forecast at time t for time t-1 (backcast)	t (nowcast)	t+1
0 months	-1 to k	0 to k	1 to k
1 month	0 to k	1 to k	2 to k
2 months	1 to k	2 to k	3 to k

Annotations: Lag -1 refers to the next month. Positive values refer to preceding months.
k: maximum lag length.

Often cross correlations between the dependent variable and the macroeconomic indicators are used to identify variables with the desired leading indicator properties (cf. e.g. Döpke et al., 1994, de Bondt and Hahn, 2010). Correlation analysis can certainly provide some useful insight, but it also has some serious weaknesses. For instance the analysis is impeded by high autocorrelation and there are distortions due to overlapping oscillations of different phase lengths (Fritsche and Stephan, 2002, p. 294).

For these reasons we have chosen a different approach. In what follows bivariate equations are estimated. The growth rate of industrial production is explained by its own lags and one indicator, the lag structure depending on the forecast horizon and the publication lags of the respective variables (cf. Table 2). All estimations are based on stationary variables. This means that in some cases variables were transformed to ensure stationarity. The transformation of data is stated in Tables 3-5. The specification of the bivariate equations is largely automated. It starts from a specification including both variables with a maximum lag length of 12. In the next step all regressors with p-values above the significance level of 0.05 are eliminated from the equation. Due to the fact that the first lags of the dependent variable are not available for the estimation there may be autocorrelation in the residuals. For this reason the robust standard errors of Newey and West (1987), which allow for autocorrelation and heteroskedasticity, are used. The bivariate equations are estimated for the full sample (1993m2 until 2010m9) and a shorter sample (1993m2 until 2006m11). The shorter sample serves as estimation period for the first 3-step forecast to be evaluated in terms of its out-of-sample forecast performance. Variations of the estimation period facilitate an analysis of the stability of the relationship between the industrial production variable and the respective explaining variable over time.

Subsequently, indicators are ranked on the basis of their contribution to explaining the log difference of industrial production (using the R^2 as a criterion) in the bivariate equations. Tables 3-5 show the results for the whole estimation period from 1993m2 until 2010m9. For the whole sample the bivariate analysis yields a clear result: the Composite Leading Indicator (CLI) of the OECD explains a relatively large share of the monthly changes in the logarithm of industrial production. The same is true of all three ifo indicators, whereas the ZEW's indicators perform poorly. The surveys of the European Commission (industrial confidence, production expectations) also show good results, although their performance declines with longer forecast horizons. By contrast, yield spreads gain explanatory power as the forecast horizon increases. Orders received from the domestic market and from abroad only help to explain the change in industrial production at very short forecast horizons. Some variables contribute little or nothing to an explanation of the dependent variable. They include new loans, the real effective exchange rate, the number of vacancies and the consumer price index. The Early Bird Indicator, published by Commerzbank, is also among the indicators which do not seem promising after preliminary analysis. However, several studies corroborate that this indicator shows a good forecast performance. Usually this is the case for forecast horizons of about one year, which are not analysed here (Robinzov and Wohlrabe, 2008, Hinze 2003).

Results are completely different if only the short sample (1993m2-2006m11) is analysed (cf. Tables 11-13 in the appendix). With the exception of the estimations for *t-1* (backcast) the explanatory power of all variables decreases dramatically for the other forecast periods. It is remarkable that now the Composite Leading Indicator does not explain movements in industrial production, whereas yield spreads are gaining relative importance for longer forecast horizons. The fact that the explanatory power of all variables is much smaller for the short sample than for the full sample suggests that a key prerequisite for the use of indicators is not fulfilled under real-time conditions: there seems to be no stable relationship between the indicator and the series to be forecast. The relatively better performance of the indicators in the full sample is possibly explained by the deep recession of 2008/2009, which is reflected in most macroeconomic indicators.

The selection of variables is based largely on the results for the full sample, because differences between individual variables are more pronounced. One may object to this on the grounds that we used information which would not have been available in real time. Or in other words: We might have chosen variables, which we would not have chosen under real-time conditions. This criticism is justified in principle, although it is only relevant in the case of one variable: the Composite Leading Indicator. By and large, despite the very different results in terms of R^2, the group of indicators selected on the basis of the full sample largely coincides with the group of indicators selected on the basis of the short sample. Whenever additional variables perform reasonably well only for the short sample, they are included in subsequent analyses. This problem is also mitigated to some extent by the fact that we take another aspect into account when selecting suitable variables: correlation between independent variables.

Table 3: Bivariate estimations for *t-1* (backcast), sample: 1993m2-2010m9

Indicator	Transformation of the indicator series	Lags of the dependent variable	Lags of the indicator series	R^2
ORDERS_DOM	log difference	1,2,5	1-4,9-10	0.37
CLI	log difference	1,3,5	3,6,9	0.37
INDCONF	level	1,3,5	1,3,6,8	0.34
PRODEX	level	1,3,5	0,7,8	0.31
IFO_EXP	log level	1,3,5	1,7	0.31
ECSENT	log level	1,3,5	0,3	0.31
EURIBOR	difference	1,3,5	-1,0,2,10	0.31
IFO_SIT	log level	1,3,5	0,3	0.30
IFO_CLIMATE	log level	1,3,5	0,7	0.29
ORDERS_FOR	log difference	1,2,4,5	1-6	0.29
YC1Y	level	1,3,5	3,5	0.27
C_SPREAD	difference	1-3,5	2,3,12	0.24
YC10Y	level	1-3,5	-1,1-4	0.24
YC2Y	level	1,3,5	3,5	0.23
OIL	log difference	1,3	0-4	0.21
YC3Y	level	1,3,5	3,5	0.20
FIX	log difference	1,2,3,5	1,3,12	0.20
YC5Y	level	1,2,3,5	3,5	0.19
ZEW_SENT	level	2,3,5	1-3	0.17
ZEW_SIT	level	1,3,5	-1,4	0.16
CDAX	log difference	2,3,5	0	0.16
CPI	log difference	2,3,5	1	0.16
VAC	log difference	2,3,5	0	0.15
REER	log difference	2,3,5	9	0.15
CREDIT	difference	2,3,5	none	0.12
COM_EB	difference	2,3,5	none	0.12

Annotations: Lag -1 refers to the next month. Positive values refer to preceding months.

Table 4: Bivariate estimations for *t* (nowcast), sample: 1993m2-2010m9

Indicator	Transformation of the indicator series	Lags of the dependent variable	Lags of the indicator series	R^2
CLI	log difference	3,5	3,6,9	0.28
IFO_SIT	log level	4,5	0,2,5,6,12	0.25
IFO_CLIMATE	log level	3,5	0,2,5,6	0.24
EURIBOR	difference	3,5	0,2,3,10	0.24
ORDERS_DOM	log difference	4,5	2,3,4,9,10	0.24
PRODEX	level	3,5	1,4	0.23
YC1Y	level	2,3,5	3,5	0.23
INDCONF	level	3,5	1,3	0.23
IFO_EXP	log level	3,5	1,2	0.22
YC2Y	level	2,3,5	3,4	0.21
YC10Y	level	3,5	0,1,2,3,4	0.21
ECSENT	log level	3,5	1,3	0.20
YC3Y	level	2,3,5	3,4	0.20
ORDERS_FOR	log difference	4,5	2,3,4,6	0.19
FIX	log difference	2,3,5	1,3,12	0.19
C_SPREAD	difference	3,5	3,12	0.18
YC5Y	level	2,3,5	3,4	0.18
ZEW_SENT	level	2,3,5	1,2,3	0.17
CPI	log difference	2,3,5	1	0.16
ZEW_SIT	level	3,5	1,4	0.16
OIL	log difference	2,3,5	0	0.15
REER	log difference	2,3,5	9	0.15
CDAX	log difference	3,5	0	0.14
VAC	log difference	3,5	0	0.13
COM_EB	difference	2,3,5	none	0.12
CREDIT	difference	2,3,5	none	0.12

Table 5: Bivariate estimations for *t+1*, sample: 1993m2-2010m9

Indicator	Transformation of the indicator series	Lags of the dependent variable	Lags of the indicator series	R^2
CLI	log difference	3,5	3,6,9	0.28
YC1Y	level	3,5	3,5	0.22
IFO_EXP	log level	3,5	1,2	0.22
IFO_CLIMATE	log level	4,5	1,5,6	0.21
YC2Y	level	3,5	3,4	0.20
IFO_SIT	log level	3,5	1,2,4-6	0.20
PRODEX	level	3,5	2,4	0.19
INDCONF	level	3,5	2,6,8	0.19
YC3Y	level	3,5	3,4	0.18
ECSENT	log level	3,5	2,3	0.17
FIX	log difference	3,5	1,3,12	0.17
C_SPREAD	difference	3	3,12	0.16
ZEW_SIT	level	3,5	1,4	0.16
YC10Y	level	3,5	2,3,4	0.15
ZEW_SENT	level	3,5	1,2,3	0.15
YC5Y	level	3	3,4	0.15
ORDERS_FOR	log difference	4,5	3,4,5,6	0.14
CDAX	log difference	3,5	2	0.12
OIL	log difference	3	3	0.10
EURIBOR	difference	3,5	none	0.09
CPI	log difference	3,5	none	0.09
ORDERS_DOM	log difference	3,5	none	0.09
COM_EB	difference	3,5	none	0.09
REER	log difference	3,5	none	0.09
CREDIT	difference	3,5	none	0.09
VAC	log difference	3,5	none	0.09

3.2 Analysis of correlation between indicators

In part there is a very high contemporaneous correlation between the individual variables (cf. Tables 6-8). This is particularly true in the case of the survey data. For instance the European Commissions's surveys – the Economic Sentiment Indicator (ECSENT), the Industrial Confidence Indicator (INDCONF) and the production expectations in industry – are highly correlated with the ifo Institute's indicators. The ZEW indicator for the economic situation is highly correlated with the respective ifo indicator as well as some series of the European Commission (ECSENT, INDCONF). This is not surprising as both the surveys of the European Commission and those of the ifo Institute refer to a similar sample of companies. Further, some series are composed of several subseries. For instance the ifo Institute's business climate index is composed of the series for the business situation and business expectations. For this reason the aggregate series and its subseries should not be used in an equation simultaneously. Similar can be said about the Economic Sentiment Indicator, the Industrial Confidence Indicator and the production expectations in industry. The latter is a subseries of the Industrial Confidence Indicator, which in turn is included in the Economic Sentiment Indicator with a weight of 40 %. The OECD's Composite Leading Indicator includes several subseries of the ifo Institute's business climate, new orders, finished goods stocks and an interest rate spread. It is striking that the correlation between the Composite Leading Indicator and its subseries which are used here is not as high as in the cases described above. Among the financial market indicators only the yield spreads show high bivariate correlations with the correlation between the yield spread of 10-year government bonds and 1-year government bonds being the lowest.

To avoid problems of multicollinearity in the multivariate estimations, indicators which show a high bivariate correlation (> 0.75 in absolute terms) are not used in the same equation. These values are in bold type in Tables 6-8. Using Table 3 as an example we explain the variable selection procedure. Actually, we go through the list from top to bottom taking the correlations between individual indicators into account. In this example, orders received from the domestic market, the CLI and the Industrial Confidence Indicator (INDCONF) are chosen initially. Production expectations in industry (PRODEX) are skipped due to their correlation with INDCONF. The next indicator to be chosen is the ifo series of business expectations, whereas the Economic Sentiment Indicator (ECSENT) is skipped due to its correlation with INDCONF. This way a list of indicators is derived. One problem consists in the fact that the list is path-dependent, i.e. with the decision in favour of INDCONF we decide against PRODEX and ECSENT, but also against the ifo Institute's business climate index and business situation. To avoid this problem, we derive several lists of indicators by e.g. choosing PRODEX instead of INDCONF, which leads to other indicators being selected in subsequent steps. The final selection among these alternatives results from a comparison of alternative specifications, which is explained in the next section.

Table 6: Selected correlations between leading indicators

	IFO_CLIMATE	IFO_SIT	IFO_EXP	ECSENT	INDCONF	PRODEX	ZEW_SENT	ZEW_SIT
IFO_CLIMATE								
IFO_SIT	**0.92**							
IFO_EXP	**0.82**	0.53						
ECSENT	**0.86**	**0.79**	0.71					
INDCONF	**0.93**	**0.89**	0.72	**0.88**				
PRODEX	**0.85**	0.64	**0.91**	**0.81**	**0.87**			
ZEW_SENT	0.12	-0.21	0.56	0.18	0.03	0.37		
ZEW_SIT	**0.81**	**0.90**	0.43	**0.81**	**0.80**	0.55	-0.25	
CLI	0.56	-0.33	0.38	-0.16	-0.25	0.09	0.60	0.38

Table 7: Selected correlations between the CLI and its subseries

	CLI
ORDERS_DOM	0.07
ORDERS_FOR	0.37
YC1Y	0.45
YC2Y	0.45
YC3Y	0.45
YC5Y	0.48
YC10Y	0.42

Table 8: Selected correlations between financial indicators

	FIX	C_SPREAD	CDAX	EURIBOR	YC1Y	YC2Y	YC3Y	YC5Y
FIX								
C_SPREAD	0.23							
CDAX	-0.25	-0.24						
EURIBOR	-0.25	0.07	0.07					
YC1Y	-0.29	-0.24	0.22	0.50				
YC2Y	-0.28	-0.24	0.20	0.40	**0.95**			
YC3Y	-0.25	-0.23	0.20	0.31	**0.88**	**0.98**		
YC5Y	-0.19	-0.68	0.19	0.19	**0.77**	**0.91**	**0.97**	
YC10Y	-0.12	-0.21	0.17	0.06	0.60	**0.79**	**0.88**	**0.97**

4 Multivariate forecast equations

In what follows multivariate linear equations are estimated for the three forecast horizons. The dependent variable is the growth rate of industrial production approximated by the log difference. It is explained by its own lagged observations and selected indicators, the lag structure depending on the forecast horizon and the publication lags of the variables used (cf. Table 2). All equations are estimated with stationary time series using OLS. This means that variables are transformed if necessary to ensure stationarity. The transformations of the data are reported in Tables 3-5. Due to the fact that, under real-time conditions, the first lags of the dependent variable are not available for estimations for period t and beyond, there may be autocorrelation in the residuals. For this reason, as in the bivariate case, the standard errors of Newey and West (1987) are used. The multivariate forecast equations are estimated for the short sample (1993m2 until 2006m11), leaving a sufficiently large number of observations for the out-of-sample forecast evaluation.

The specification of the multivariate equations is based on a largely automated general-to-specific approach. As we permit a maximum lag length of 12 for all variables in the initial equation, there are restrictions to the number of variables we can include without encountering problems with the number of degrees of freedom. Therefore, the number of indicators is limited to eight. Insignificant regressors are eliminated in a procedure similar to the one described in Section 3.1. Regressors which show a p-value above a pre-defined threshold are eliminated. This is done in several steps beginning with a threshold of 0.4, which is gradually lowered to 0.05. On the one hand the automatisation has the advantage that a large number of equations can be specified and evaluated. On the other hand it leads to additional restrictions. For instance, insignificant variables have to be eliminated in a sequential procedure. This means that the algorithm is applied to one variable after another, while all other variables are ignored. When the selection process is completed for the first variable the programme proceeds to the next. This approach carries the risk that the result of variable selection depends on the order of indicators in the elimination process. We take this problem into account by arranging the order of the variables in a way that the procedure starts with the weakest variable, whereas the more promising variables are tested later. In addition we partly altered the order of the variables in the selection process. The resulting variable selection turned out very similar in all cases.

As described in Section 3.1 the list of indicators is path-dependent and there are several lists for each forecast horizon. Which of the alternatives is selected as the "final" forecast equation is determined in a comparison of the specifications. The following criteria are used:

- the fit of the estimation equation in different estimation periods (in terms of the R^2),
- the evaluation of the in-sample forecast performance (in terms of the root mean squared error and Theil's inequality coefficient – especially its variance and bias proportion), and
- whether the residuals of the estimation equation are normally distributed.

For each forecast horizon the specification which performs best according to the criteria mentioned above is chosen. Table 9 provides an overview of the regressors which are relevant for each forecast horizon. In each case a combination of "hard" and "soft" indicators was selected in addition to the autoregressive term. In the case of the backcast (genuine forecast for period *t-1*) the autoregressive term plays an important part absorbing the autocorrelation. With an increasing forecast horizon its importance decreases, which is due to the fact that the first few lags are not available due to the publication lag. Table 9 largely confirms the results of the preliminary bivariate analysis. Nevertheless, there are additional insights.

Table 9: Regressors for individual forecast horizons

	t-1	*t*	*t+1*
Dependent variable	IPRO	IPRO	IPRO
Real economic indicators	ORDERS_FOR	ORDERS_FOR	ORDERS_FOR
	ORDERS_DOM	ORDERS_DOM	
Surveys		IFO_CLIMATE	IFO_CLIMATE
	IFO_EXP		
	IFO_SIT		
		ZEW_SENT	ZEW_SENT
Composite indicators	CLI	CLI	CLI
Financial indicators	EURIBOR	EURIBOR	
	YC1Y		YC1Y
		YC10Y	YC10Y
	FIX	FIX	FIX
			CRP_SPRD

It is remarkable that orders received from abroad play a role for all three forecast horizons, although, in the bivariate analysis, they were only observed to perform well in the very short term. Among the survey indicators the ifo Institute's indicators perform well as expected. However, it is surprising that the ZEW's economic sentiment indicator, which failed in the preliminary analysis, is useful in combination with other indicators. The CLI meets the expectations we had after the preliminary analysis. By contrast, the performance of the financial market indicators is very remarkable. On the one hand, the result of the preliminary analysis, that the yield spread gains importance with an increasing forecast horizon, is confirmed. On the other hand, it is surprising that both the corporate spread and the fixed income index have explanatory potential when combined with other variables.

5 Estimation of the recession probability

As mentioned above, the initial forecast equations are based on the estimation sample 1993m2 until 2006m11; this is the data base which is available at publication period 2007m1. On this basis, we predict the log differences of industrial production in 2006m12, 2007m1 and 2007m2. With each new release of the industrial production index the estimation sample is increased by one observation and the three forecast equations are re-estimated. This means that the specifications – especially the particular lag structures – change each time. The forecasting exercise is carried out for each publication period within the evaluation period (2007m1-2010m11) leading to 47 sets of predictions.

The predictions of the growth rates of industrial production are merely an interim result. Our focus is on the recession probability at each of the 47 publication periods (= vintages). Therefore, we calculate for each vintage cumulative growth rates over five months (the defined minimum duration of a recession) and identify on the basis of the density function of the cumulative growth rate the probability that the estimated growth falls below a pre-defined recession threshold. If the probability is larger than 50 %, this is interpreted as a recession signal.

As no five-step-ahead forecast is carried out, the cumulative growth rate over five months is derived from a combination of published data and forecasts of the growth rates. In the first case ("1-step forecast") four ex post observations are combined with a one-step-ahead forecast (based on the equation for *t-1*). In the second case ("2-step forecast") three ex post observations are combined with one two-step-ahead forecast (based on the equations for *t-1* and *t*). In the third case ("3-step forecast") two ex post observations are combined with a three-step-ahead forecast (based on the equations for *t-1*, *t*, *t+1*).

The density function of the cumulative growth rate can be derived as follows: From the normal distribution of the individual forecast errors we can derive the normal distribution of the cumulative forecast error over five months. Furthermore, the variance of the sum of residuals is equal to the sum of all elements of the covariance matrix. With its mean and variance the normal distribution is uniquely determined. Hence, the density function of the cumulative growth rate is determined, too. Note, that we have a specific density function for each single cumulative growth rate at each publication period. This is due to the fact that with each new release of the industrial production index the three forecast equations are re-estimated implying that the corresponding residual vectors change at each time. This means that the respective covariance matrices change, too. Given the specific density function of each cumulative growth rate we can identify the probability that the estimated growth falls below the pre-defined recession threshold.

The determination of the threshold is based on the triangular approximation approach of Harding and Pagan (2002) which uses the duration and the amplitude of the cycle as key criteria in the determination of turning points. A recession is indicated when the area of the triangle determined by the duration and amplitude falls below a certain threshold. In what follows we apply the decision rule that (a) a recession must have a minimum duration of five months and (b) the output variable must decline by at least 1 % compared to the local maximum during these five months. Based on these parameters we obtain a threshold of -0.025[1].

6 Forecast evaluation under real-time conditions

In what follows the three forecast equations are evaluated under real-time conditions. In their assessment the key issue is whether they indicated the beginning and the end of a recession at an earlier point in time (with respect to the publication period) than official statistics would have permitted. As a benchmark we use a binary series which assumes the value 0 in expansions and 1 in contractions. This series was generated by

[1] $0.5 * (a) * (b) <$ threshold, with $(a) = 5$ and $(b) = -1$.

means of a non-parametric dating algorithm, which is based on the procedure of the National Bureau of Economic Research (NBER) for detecting turning points (Bry and Boschan, 1971). In addition, a minimum amplitude adapted to phases of low volatility as proposed in the triangular approximation approach of Harding and Pagan (2002) is required.[2] This sophisticated dating procedure corresponds to the current state of research. However, it is only suitable for ex post business cycle dating, because, to determine whether a local extremum is also a turning point, it requires both past and future observations which are not available under real-time conditions.

For each vintage (out of a total of 47) Figure 1 shows the recession probabilities which were estimated on the basis of the one-step, the two-step and the three-step-ahead forecast. The grey area represents a phase during the evaluation period for which the benchmark is equal to one thus indicating a recession. It is important to note that this series, too, refers to the publication period. This means it does not reflect in which month a turning point occurred, but when the latter could be reported for the first time on the basis of published data. Due to the defined minimum duration of a recession of five months and the publication lag of two months in the case of industrial production a recession signal can occur seven months after the turning point at the earliest. For this reason the benchmark series had to be shifted right on the time axis by seven months. Good forecasts would indicate a recession ahead of the benchmark, i.e. the recession probability should rise above the threshold of 50 % already before (i.e. left of) the benchmark.

As Figure 1 illustrates, official data published in October 2008 provided an ex post signal of a recession. It is remarkable that all three forecast models already provide the signal at an earlier point in time. In the case of the 1-step and the 2-step forecast the lead is two months, whereas it is only one months in the case of the 3-step forecast.[3] The forecast equations are even better at predicting the end of a recession: Whereas, based on the published data, the end of the recession could not be confirmed before December 2009, the 3-step forecast signalled expansion as early as in July 2009, i.e. five months earlier. Both the 1-step and the 2-step forecast indicated in August 2009 that the recession was over and thus still had a time advantage of four months over the ex post dating based on official data.

[2]Cf. Schreiber et al., 2012, Section 4.3.2.

[3]However, the recession probability forecast in August 2008 on the basis of the 3-step forecast equalled 49 % and was thus only slightly below the recession threshold of 50 %.

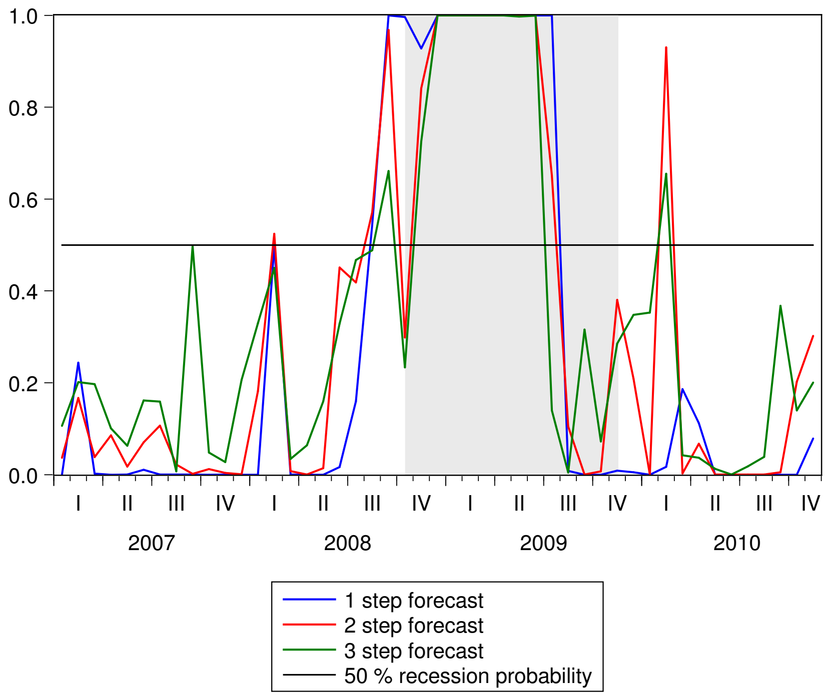

Figure 1: Recession signals related to the publication period

In addition to an early detection of recession phases, correct signals are a key evaluation criterion. All forecasts fulfil this criterion in the sense that they do not miss a recession during the evaluation period, which would have been nearly impossible in the light of the massive recession of 2008/2009. By contrast, a stricter criterion is whether the forecasts gave incorrect signals, i.e. whether a recession was erroneously reported. Figure 1 shows that both the 2-step and the 3-step forecast provided two clearly wrong signals during the evaluation period: in October 2008 the incorrect signal "no recession" was given during the recession phase, whereas in March 2010 the erroneous signal "recession" was given during an expansion. Basically, the problem of wrong signals can be alleviated by pooling the 1-step, the 2-step and the 3-step forecast for the same month. Figure 1 illustrates that the volatility of the forecast recession probability increases significantly as the forecast horizon increases. This is not surprising. As in the case of the 3-step forecast two ex post observations are combined with a three-step-ahead forecast whereas in the case of the 1-step forecast four ex post observations are combined with a one-step-ahead forecast the uncertainty and thus the volatility is clearly higher in the former than in the latter case. In terms of the chosen evaluation criteria the predicted recession probabilities based on the 1-step forecast perform best.

7 Summary

In this paper we describe an attempt to develop a methodology which can help to improve the assessment of the current economic situation. Whereas most of the relevant studies predict the development of the reference series (which reflects the movements of the business cycle), we estimate the recession probability. We propose an approach which combines multivariate single equations to forecast the monthly growth rate of industrial production with a density forecast. On this basis we can derive the current recession probability. In the analysis the focus was on the real-time problem, i.e. the fact that the reference series as well as important indicators are not available on a timely basis and are often revised substantially over an extended period. For this reason the whole analysis (estimation of forecast equations, forecast of the monthly growth rates of industrial production, evaluation of the estimated recession probabilities) was carried out under real-time conditions.

The results presented here constitute the first step of a long journey. Although they are encouraging, there is still a substantial need for further research. Two particularly urgent problems should be mentioned here: During the analysed evaluation period the estimated recession probabilities allowed to identify the beginning and the end of the recession in 2008/2009 ahead of official data. Whereas the beginning of the recession was detected with a lead of 1-2 months compared to the business cycle dating procedure based on official data, the end was forecast with a much longer lead of 4-5 months. This is irritating, because the maximum forecast horizon is only 3 months. This result is a consequence of the chosen benchmark and, once again, raises the question which benchmark should be used in the evaluation of the forecast equations. As the whole analysis was undertaken under strict real-time conditions, it would have been desirable to use a real-time benchmark. However, as explained in detail, this was not possible, because for the assessment of whether there is a turning point at time t the dating algorithm needs both past and future observations of industrial production. After extensive considerations which benchmark to use in the assessment we decided in favour of the business cycle dating based on the final vintage (November 2010). However, this implies that the revision process is already completed for the observations which are relevant for the dating of the recession of 2008/2009. This means that the benchmark reflects the business cycle as it is seen long after the initial publication period. Figure 2 illustrates the relevance of this issue showing the revision history of the binary benchmark series as it is recorded in the upper triangular matrix (UTM). The UTM is the usual way real-time data are presented, because it displays the two relevant dimensions of real-time data: the *reporting* period (row) and the *publication* period (column). The reporting period refers to the point in time when the economic activity actually took place, for example the industrial production in January 2005. The publication period, however, refers to the point in time when the data concerning the economic activity are published. Due to the publication lag of two months data for industrial production in January 2005 were published for the first time in March 2005. In the consecutive months the data for industrial production in January 2005 were revised repeatedly due to late registrations, error corrections etc. until the revision process was completed. For each reporting period the revision process of the variable is recorded in the corresponding row of the UTM.

reporting period (row)	publication period (column) 2009M3	2009M4	2009M5	2009M6	2009M7	2009M8	2009M9	2009M10	2009M11	2009M12	2010M1	2010M2
2008M01	0	0	0	0	0	0	0	0	0	0	0	0
2008M02	0	0	0	0	0	0	0	0	0	0	0	0
2008M03	0	1	1	1	1	1	1	1	1	1	1	1
2008M04	0	1	1	1	1	1	1	1	1	1	1	1
2008M05	0	1	1	1	1	1	1	1	1	1	1	1
2008M06	0	1	1	1	1	1	1	1	1	1	1	1
2008M07	0	0	1	1	1	1	1	1	1	1	1	1
2008M08	0	0	0	1	1	1	1	1	1	1	1	1
2008M09	0	0	0	0	1	1	1	1	1	1	1	1
2008M10		0	0	0	0	1	1	1	1	1	1	1
2008M11			0	0	0	0	1	1	1	1	1	1
2008M12				0	0	0	0	1	1	1	1	1
2009M01					0	0	0	0	1	1	1	1
2009M02						0	0	0	0	1	1	1
2009M03							0	0	0	0	1	1
2009M04								0	0	0	0	1
2009M05									0	0	0	0
2009M06										0	0	0
2009M07											0	0
2009M08												0

Figure 2: Industrial production: revision history for the recession period 2008/2009

Figures 2 shows that the revisions concerning the data for industrial production obviously changed the assessment of the dating algorithm concerning the beginning of the recession. On the basis of the official data published in March 2009 no recession is indicated after March 2008. One month later, however, the recession becomes apparent in the official data. Further, it is remarkable that in each subsequent month one month is added to the recession phase. Figure 2 shows that the revisions of the industrial production data make an assessment of the current economic situation extremely difficult or even impossible. Therefore, it is necessary to use indicators for this purpose as they are revised to a lesser extent or not at all. Obviously, the cyclical tendency already becomes apparent in the indicators at a time when it cannot yet be seen in the industrial production data. It might therefore be helpful to include a series of first revision of industrial production as an additional regressor in the forecast equation.

The second important issue is that recessions are rare events. Even in long time spans they do not occur often. In our evaluation period there is only one recession. Therefore, the estimation period of our forecast equations should be shortened, so that there will be more than one recession in the evaluation period. Further, the forecast equations should also be estimated with a moving window of a fixed number of observations to examine what will happen, if we shorten the "learning phase".

8 References

Bry, G. and Boschan, C. (1971), *Cyclical analysis of time series: Selected procedures and computer programs*, National Bureau of Economic Research, Inc.

Carstensen, K., Wohlrabe, K., Ziegler, C. (2011), "Predictive ability of business cycle indicators under test - a case study for the euro area industrial production", *Jahrbücher für Nationalökonomie und Statistik* **231** (1), 82-106.

de Bondt, G. and Hahn, E. (2010), "Predicting recessions and recoveries in real time. The euro area-wide leading indicator (ALI)", *ECB Working Paper* (1246).

Dickey, D.A. and Fuller, W.A. (1979), "Distribution of the estimators for autoregressive time series with a unit root", *Journal of the American Statistical Association* **74**, 427-431.

Dickey, D.A. and Fuller, W.A. (1981), "Likelihood ratio statistics for autoregressive time series with a unit root", *Econometrica* **49**, 1057-1072.

Döpke, J., Krämer, J.W., Langfeldt, E. (1994), "Konjunkturelle Frühindikatoren in Deutschland", *Applied Economics Quarterly (Konjunkturpolitik)* **40**, 135-153.

Drechsel, K. and Scheufele, R. (2010), "Should we trust in leading indicators? Evidence from the recent recession", *IWH Discussion Paper* (10-10).

Drechsel, K. and Scheufele, R. (2011), "The performance of short-term forecasts of the German economy before and during the 2008/2009 recession", *International Journal of Forecasting* (forthcoming)

Forni, M., Hallin, M., Lippi, M., Reichlin, L. (1999), "The generalized dynamic factor model: Identification and estimation, *CEPR Discussion Paper* (2338)

Fritsche, U. and Stephan, S. (2002), "Leading indicators of German business cycles. An Assessment of properties", *Jahrbücher für Nationalökonomie und Statistik* **222** (3), 289-315.

Fuller, W.A. (1976), *Introduction to statistical time series*, Wiley, New York.

Granger, C. (1969), "Investigating causal relations by econometric models and cross-spectral methods", *Econometrica* **37** (3), 424-438.

Harding, D. and Pagan, A. (2002), Dissecting the cycle: A methodological investigation", *Journal of Monetary Economics* **49** (2), 365-381.

Hinze, J. (2003), "Prognoseleistung von Frühindikatoren. Die Bedeutung von Frühindikatoren für Konjunkturprognosen - eine Analyse für Deutschland", *HWWA Discussion Paper* (236).

Newey, W. and West, K (1987), "A simple, positive, semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix", *Econometrica* **55** (3), 703-708.

Robinzonov, N. and Wohlrabe, K. (2008), "Freedom of choice in macroeconomic forecasting: An illustration with German industrial production and linear models", *Ifo Working Paper* (57).

Schreiber, S., Theobald, T., Proano, C., Stephan, S., Rietzler, K., Detzer, D. (2012), *Verfahren der konjunkturellen Wendepunktbestimmung unter Berücksichtigung der Echtzeit-Problematik. Welche Indikatoren geben frühzeitig zuverlässige Signale?*, Gutachten im Auftrag des Bundesministeriums der Finanzen, *IMK Study* Nr. 27.

Schumacher, C. and Breitung, J. (2008), "Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data", *International Journal of Forecasting* **24** (3), 386-398.

Stark, T. and Croushore, D. (2002), "Forecasting with a real-time data set for macroeconomists", *Journal of Macroeconomics* **24** (4), 507-531.

9 Appendix

9.1 Data

Table 10: Results of the ADF-Tests

Variable	**Transformation**	**Deterministic**	**Lags**	**Test statistics**	**Results**
IPRO	log level	constant, trend	1-3	-3.22*	I(1)
	log difference	constant	1-2	-5.74***	I(0)
ORDERS_DOM	log level	constant, trend	1-2	-2.90	I(1)
	log difference	constant	1	-8.77***	I(0)
ORDERS_FOR	log level	constant, trend	1-3	-2.59	I(1)
	log difference	constant	1-2	6.63***	I(0)
VAC	log level	constant	1-3	-2.35	I(1)
	log difference		1-2	-3.92***	I(0)
IFO_CLIMATE	log level	constant	1-3	-3.80***	I(0)
	log difference		1-2	-4.42***	I(0)
IFO_SIT	log level	constant	1-5	-4.15***	I(0)
	log difference		1-2	-4.39***	I(0)
IFO_EXP	log level	constant	1-3	-4.45***	I(0)
	log difference		0	-10.18***	I(0)
ZEW_SENT	level	constant	1	-4.07***	I(0)
	difference		0	-8.23***	I(0)
ZEW_SIT	level	constant	1-3	-3.42**	I(0)
	difference		1	-4.73***	I(0)
INDCONF	level	constant	1-3	-4.51***	I(0)
	difference		1	-4.70***	I(0)
PRODEX	level	constant	1-2	-3.82***	I(0)
	difference		1	-6.62***	I(0)
ECSENT	log level	constant	1-2	-3.43**	I(0)
	log difference		1	-5.43***	I(0)
CLI	log level	constant, trend	1-6	-4.12***	trend stat.
	log difference	constant	1-7	-4.68***	I(0)
COM_EB	level	constant	0	-1.66	I(1)
	difference		0	-14.02***	I(0)
CPI	log level	constant, trend	0	-2.70	I(1)
	log difference	constant	0	-15.27***	I(0)
REER	log level	constant	1	-1.69	I(1)
	log difference		0	-11.41***	I(0)
OIL	log level	constant, trend	1	-3.12	I(1)
	log difference	constant	0	-12.11***	I(0)
CDAX	log level	constant	0	-2.00	I(1)
	log difference		0	-13.51***	I(0)
CREDIT	level	constant	0	-1.94	I(1)
	difference		0	-10.57***	I(0)
FIX	log level	constant, trend	0	-2.42	I(1)
	log difference	constant	0	-13.06***	I(0)
EURIBOR	level	constant	1	-2.48	I(1)
	difference		0	-7.81***	I(0)
C_SPREAD	level	constant	0	-2.20	I(1)
	difference		0	-13.47***	I(0)

Variable	Transformation	Deterministic	Lags	Test statistics	Results
YC1Y	level	constant	1	-4.27***	I(0)
	difference		0	-11.93***	I(0)
YC2Y	level	constant	1	-3.99***	I(0)
	difference		0	-12.24***	I(0)
YC3Y	level	constant	1	-3.70***	I(0)
	difference		0	-11.85***	I(0)
YC5Y	level	constant	1	-3.26**	I(0)
	difference		0	-11.25***	I(0)
YC10Y	level	constant	1	-2.73*	I(1)
	difference		0	-11.06***	I(0)

***, **, * denote significance at 1%, 5%, and 10% level.

9.2 Bivariate estimations

Table 11: Bivariate estimations for *t-1* (backcast), sample: 1993m2-2006m11

Indicator	Transformation of the indicator series	Lags of the dependent variable	Lags of the indicator series	R^2
IFO_EXP	log level	1,2	1,3,6,7,10,11	0.33
EURIBOR	difference	1,2	0,2,11	0.32
ECSENT	log level	1,2	0,3,6,7	0.30
IFO_SIT	log level	1,2,11,12	0,4,11,12	0.30
ORDERS_FOR	log difference	1,2,3	1-6	0.30
IFO_CLIMATE	log level	1,2	1,2,11	0.29
YC10Y	level	1,2	0-4	0.29
CDAX	log difference	1,2,11,12	-1,0,4,5,7	0.29
CLI	log difference	1,2	2	0.28
INDCONF	level	1,2	0,7,8	0.28
FIX	log difference	1,2	3,9	0.28
PRODEX	level	1,2	0,7	0.27
YC1Y	level	1,2	3,4	0.27
YC2Y	level	1,2	3,4	0.27
ZEW_SENT	level	1,2	-1,4	0.27
YC3Y	level	1,2	3,4	0.26
ZEW_SIT	level	1,2	-1,5	0.26
YC5Y	level	1,2	3,4	0.24
VAC	log difference	1,2	0	0.22
REER	log difference	1,2	3,9	0.21
CPI	log difference	1,2	7	0.20
C_SPREAD	difference	1,2	3,12	0.19
ORDERS_DOM	log difference	1,2	none	0.18
OIL	log difference	1,2	none	0.18
CREDIT	difference	1,2	none	0.18
COM_EB	difference	1,2	none	0.18

Lag -1 refers to the next month. Positive values refer to preceding months.

Table 12: Bivariate estimations for t (nowcast), sample: 1993m2-2006m11

Indicator	Transformation of the indicator series	Lags of the dependent variable	Lags of the indicator series	R^2
YC10Y	level	3	0-4	0.17
YC5Y	level	none	2-4	0.11
YC1Y	level	none	3,4	0.09
YC2Y	level	none	3,4	0.09
YC3Y	level	none	3,4	0.09
FIX	log difference	none	3,9	0.09
IFO_SIT	log level	none	0,2,5,6	0.07
IFO_CLIMATE	log level	none	0,2,5,6	0.07
ORDERS_FOR	log difference	none	3,6	0.07
EURIBOR	difference	none	2,3	0.06
ORDERS_DOM	log difference	3,11	6,9	0.06
CDAX	log difference	3	0	0.05
CPI	log difference	3	7	0.04
CLI	log difference	none	2	0.03
C_SPREAD	difference	none	6	0.03
REER	log difference	none	3	0.03
COM_EB	difference	none	10	0.02
VAC	log difference	none	0	0.01
PRODEX	level	none	none	0.00
INDCONF	level	none	none	0.00
IFO_EXP	log level	none	none	0.00
ECSENT	log level	none	none	0.00
ZEW_SENT	level	none	none	0.00
ZEW_SIT	level	none	none	0.00
OIL	log difference	none	none	0.00
CREDIT	difference	none	none	0.00

Table 13: Bivariate estimations for t+1, sample: 1993m2-2006m11

Indicator	Transformation of the indicator series	Lags of the dependent variable	Lags of the indicator series	R^2
YC10Y	level	3	2-4	0.13
YC5Y	level	none	2-4	0.11
YC1Y	level	none	3,4	0.09
YC2Y	level	none	3,4	0.09
YC3Y	level	none	3,4	0.09
FIX	log difference	none	3,9	0.09
IFO_CLIMATE	log level	11	1,2,3,5,6	0.07
IFO_SIT	log level	11	5,6,10,12	0.07
ORDERS_FOR	log difference	none	3,6	0.07
EURIBOR	difference	none	2,3	0.06
CPI	log difference	3	7	0.04
C_SPREAD	difference	none	6	0.03
REER	log difference	none	3	0.03
ECSENT	log level	none	2,3	0.02
COM_EB	difference	none	10	0.02
CLI	log difference	none	none	0.00
IFO_EXP	log level	none	none	0.00
PRODEX	level	none	none	0.00
INDCONF	level	none	none	0.00
ZEW_SIT	level	none	none	0.00
ZEW_SENT	level	none	none	0.00
CDAX	log difference	none	none	0.00
OIL	log difference	none	none	0.00
ORDERS_DOM	log difference	none	none	0.00
CREDIT	difference	none	none	0.00
VAC	log difference	none	none	0.00

Publisher: Hans-Böckler-Stiftung, Hans-Böckler-Str. 39, 40476 Düsseldorf, Germany
Phone: +49-211-7778-331, IMK@boeckler.de, http://www.imk-boeckler.de

IMK Working Paper is an online publication series available at:
http://www.boeckler.de/imk_5016.htm

ISSN: 1861-2199

Hans Böckler Stiftung

Fakten für eine faire Arbeitswelt.