Diskussionspapiere
Discussion Papers

Discussion Paper No. 207

Leading Indicators of German Business Cycles: An Assessment of Properties

by
Ulrich Fritsche and Sabine Stephan

Berlin, April 2000

Deutsches Institut für Wirtschaftsforschung, Berlin
Königin-Luise-Str. 5, 14195 Berlin
Phone: +49-30-89789- 0
Fax: +49-30-89789- 200
Internet: http://www.diw.de
ISSN 1433-0210

Leading Indicators of German Business Cycles: An Assessment of Properties*

by

Ulrich Fritsche and Sabine Stephan

Abstract:

A reliable leading indicator should possess the following properties: (1) The movements in the indicator series should resemble those in the business cycle reference series. (2) The relation between the reference series and the indicator should be statistically significant and stable over time. (3) The inclusion of the indicator in out-of-sample forecasting procedures should improve the predictive power. Our analysis deals with tests for these requirements applied to German data. We used frequency domain analysis, different Granger-causality tests and out-of sample forecasts. Only few indicators passed all tests. Their inclusion into VAR-based forecasts improves the forecast in the very short run.

Konjunkturelle Frühindikatoren für Deutschland: Eine Untersuchung ihrer Eigenschaften

Brauchbare Frühindikatoren sollten folgende Eigenschaften besitzen: (1) Die konjunkturellen Bewegungen des Frühindikators sollten denen der Referenzreihe folgen. (2) Die Beziehung zwischen den Reihen sollte stabil und signifikant sein. (3) Die Einbeziehung des Indikators sollte die Out-of-sample-Prognose verbessern. Unsere Untersuchung testet diese Anforderungen für deutsche Daten. Dazu werden Methoden der Spektralanalyse, verschiedene Granger-Tests und Out-of-sample-Prognosen verwendet. Nur wenige Indikatoren bestehen die Tests auf die geforderten Eigenschaften. Ihre Einbeziehung in VAR-basierte Prognosen verbessert die Prognoseleistung in der sehr kurzen Frist.

JEL Classification: E 32, L 60, L 70

Addresses:

Ulrich Fritsche
German Institute for Economic Research (DIW)
Department of Business Cycles and Forecasting
Königin-Luise-Straße 5
D-14195 Berlin
Phone: +4930/89 789 315
e-mail: ufritsche@diw.de

Sabine Stephan
German Institute for Economic Research (DIW)
Department of Business Cycles and Forecasting
Königin-Luise-Straße 5
D-14195 Berlin
Phone: +4930/89 789 417
e-mail: sstephan@diw.de

* We would like to thank the participants of the "Seminar for Quantitative Economic Research" at the Freie Universität Berlin, especially Professor Jürgen Wolters, as well as our colleagues at the German Institute for Economic Research (DIW) in Berlin, especially Gustav-Adolf Horn, for helpful comments and fruitful discussions about the topic. We thank Jörg Döpke for sharing some of his programmes and benefitting from his ongoing research. Thanks to Jan Gottschalk for comments and to Kirsten Lommatzsch who helped to improve our English version.

1. Introduction

The application of business cycle indicators has been a means of studying and forecasting cycle movements from the beginning of business cycle research. Among all indicators, leading indicators are of special interest since they can improve the power of business cycle forecasts (especially the prediction of turning points and quantitative prognosis).[1]

A reliable leading indicator should possess the following properties:

(1) Movements in the indicator series should resemble those in the business cycle reference series.

(2) The relationship between the reference series and the indicator should be statistically significant and stable over time. Moreover, the inclusion of the indicator should improve the predictive power over that of a simple autoregressive process.

(3) The inclusion of the indicator in out-of-sample forecasting procedures should improve the predictive power (compared to a "naive" prognosis).

Our analysis deals with tests for these requirements applied to German data. First of all we have to decide which of the potential indicators selected on theoretical grounds are related to the business cycle reference series. One approach to the investigation of time series properties, which is rarely used nowadays, is frequency domain analysis. This method is used as a test for meeting the first requirement, namely the test of co-movements in the indicator and reference series. The coherence measure used in the frequency domain approach enables us to measure the strength of the relationship between the business cycle reference series and the indicator series. In accordance with the significance band proposed by Koopmans, we confined the choice of potential leading indicators to those which have a significant relationship in the relevant interval. In our further investigations, we only included those indicators which passed the first test in the frequency domain.

A reliable leading indicator should have a stable lead-structure relative to the movements of the cycle and should furthermore improve the forecasting power of cycle movements compared with simple autoregressive processes. After testing for possible lead-lag structures with cross-

---

[1] Indicators can be divided in leading, coincident and lagging indicators. Furthermore there are indicators which show the degree of "tightness" on markets. The first attempts to describe cyclical movement of economic activity based on a system of indicators date back to the early twenties. The "Harvard Barometer" - published between 1919 and 1922 - was one of the first well-known leading indicator systems. Its construction was based on 13 time series. The *Deutsches Institut für Konjunkturforschung* (later DIW) under the leadership of its first president Professor Ernst Wagemann established the first leading indicator system for Germany. The indicator-based research of Burns and Mitchell at NBER in the thirties and forties helped to establish the Anglo-Saxon view of the "business cycle as consensus". Cf. Tichy (1994); Wagemann (1928); Burns/Mitchell (1946); Oppenländer (1997); Moore/Zarnowitz (1986).

correlograms we used the criterion of Granger-causality to decide which indicators meet the second requirement of our list. Because traditional pair-wise Granger-causality tests possess some pitfalls, we carried out a modified Granger-causality test as well.[2]

Unfortunately, it is by no means certain that the indicators with the best in-sample performance perform equally well in out-of-sample forecasts. Therefore, for indicators that passed all earlier tests, we examined their forecast performance using a procedure proposed by Davis and Fagan, thus testing for the third requirement for a reliable indicator.[3]

2. The Data

2.1. Choice of Variables

For decades, the use of leading indicators in business cycle research has been criticised for being "measurement without theory".[4] There are however a number of rationales that underlie indicator choice and justify research on leading indicators. The most important rationales are *production time* (time between ordering and production); *ease of adaptation* (some aggregates are affected by short-term fluctuations earlier and/or stronger than others); *market expectations* (some series reflect or react to anticipations of future economic activity) and *prime movers* (economic fluctuations are driven by measurable economic forces such as monetary policy).[5] Furthermore indicators are often chosen for their resistance against revisions, as well as early availability. For instance, monetary indicators are available sooner than most other indicators.

In particular, indicators are of crucial importance to applied business cycle forecasting. In recent years, examination of their properties has gained considerable attention from researchers. Besides the often-cited American works of Stock and Watson (Stock/Watson 1989), there are a number of German examinations of leading indicators of business cycles. These include the articles of Döpke/Krämer/Langfeldt (1994), Langfeldt (1994), Köhler (1994), Sauer/Scheide (1995), Funke (1997), Seifert (1999) and Langmantel (1999).

In determining the business cycle reference series, we relied on a "narrow" interpretation of the business cycle and chose industrial production (excluding construction). After the introduction of the new industrial classification (NACE or WZ 93 for Germany), this series was re-estimated and prolonged by Eurostat back to 1978. This is why our analysis starts in 1978.

---

[2] Cf. Wolters (1996).

[3] Cf. Davis/Fagan (1997).

[4] This is known as "The Koopmans Critique". Cf. Koopmans (1947); Klein (1997).

[5] Cf. De Leeuw (1991).

We basically included the same indicators in our analysis as Döpke/Krämer/Langfeldt (1994). These indicators are common in the German leading indicator discussion and the results can therefore be easily compared. However for our analysis some indicators were excluded as for example, the number of *Kurzarbeiter*[6]. Our choice of indicators can be justified for several reasons. One group, the order inflows, was chosen on the grounds of production technology, since on the macroeconomic (aggregate) level we expect a relatively stable relationship between the inflow of orders and production. The choice of other indicators is justified by the fact that these indicators contain information about market expectations. In particular, this applies to the *ifo* indicators (business climate and business expectations) and the consumer confidence indicator, which are designed to measure expectations. Furthermore, we included the spread between government bond yields (assumed to carry no risk) and private bond yields (which can reflect uncertainty regarding future economic activity).[7] This measure should provide information on confidence in the economy. For a number of indicators, namely the *ifo* indicators and production indices, we used indicators which refer to the manufacturing industry, to producers of investment goods and producers of intermediate inputs. This reflects the idea that some sectors of the economy are leading or lagging compared with the overall business cycle - a view popular already in traditional business cycle theory and taken up again by real business cycle approaches[8].

The use of monetary indicators can be justified in several ways. On the one hand some business cycle theories emphasise the role of monetary developments in determining business cycle movements. In particular, this is the case in so-called "monetary over-production theories".[9] The argument that monetary developments influence business cycle movements can likewise be applied to the role of interest rates in determining economic decisions (for instance investment decisions) - especially in Keynesian business cycle theories. On the other hand, it can be assumed that all monetary indicators reflect expectations regarding the future path of economic activity.[10] As mentioned above, monetary indicators are available sooner than most other indicators.

---

[6] The main problem with this variable is that, after German reunification, this was an instrument for reducing labour volume, which was intensively used especially in Eastern Germany due to a changed incentive structure for the enterprises. Thus, this is not the "traditional behaviour" that depends on the position within the business cycle. Technically speaking, one can find strong evidence of a structural break.

[7] We calculated the difference between the *Umlaufsrendite öffentlicher Anleihen* and the *Umlaufsrendite der Industrieobligationen,* Cf. Friedman/Kuttner (1992) for theoretical arguments.

[8] Cf. Haberler (1948[2]); Entorf (1990)

[9] Cf. Hayek (1931), Haberler (1948[2]).

[10] For the monetary aggregate indicators we calculated nominal and "real" monetary aggregates, taking the contemporary consumer price index as the deflator. To calculate these measures more accurately in terms of mainstream monetary theory, "expected" inflation should be used instead of actual inflation. This is however rather difficult to measure. This argument holds both for monetary aggregates and for the calculation of real interest rates, but it is much more important in the latter case, if one considers the famous Fisher equation. One attempt to solve this problem is the approach of Mishkin (1981). Other authors calculate trend functions. In our analysis, we followed a compromise strategy. In the case of monetary aggregates we calculated "real" aggregates using the actual consumer

The real effective exchange rate was included because of a common argument which states that most booms in Germany are initiated by export-led upswings, which in turn are based on improved competitiveness.

The time series for order inflows and production were provided by Eurostat; *ifo* series for climate and expectations were calculated by the Munich based ifo-Institute. Monetary indicators, as well as interest rates were obtained from the Deutsche Bundesbank. The spread between government and private bond yields was calculated by the authors using data provided by the Deutsche Bundesbank. The consumer sentiment indicator, as well as the real effective exchange rate are from the OECD database. We chose monthly data in order to attain more accuracy in identifying turning points. Furthermore, this ensured that there were sufficient degrees of freedom available for non-parametric estimation in the frequency domain. Estimations were carried out for the period from 1978:1 to 1998:12.

The structural break caused by German reunification implies that econometric testing may face some difficulties. Eurostat, which provided the time series for the different production indices as well as for the order inflows, chained the time series for West Germany (up to 1990) and Germany (from 1991 onwards). A chaining procedure was also used for the monetary aggregates. The *ifo* indicator series are time series for West Germany only. All other time series refer to West German data until reunification and to German data afterwards.

2.2. Data Properties

Most of our procedures require stationarity assumptions. Therefore we tested all time series for unit roots using augmented Dickey-Fuller tests.[11] Non-stationary variables were transformed into stationary variables by calculating annual growth rates. This strategy has the advantage that highly complicated de-trending procedures are avoided, the results of which depend on the assumed structure of the data generating process.[12] Furthermore, as some studies have shown, the chosen filtering procedure has the advantage that a lot of spectral density remains in the region relevant for our topic.[13] In addition, annual growth rates are quite often used for forecasting purposes and economic policymaking in Germany. Furthermore, the transformation into annual growth rates serves as a simple method of seasonal adjustment. All stationary variables that remained specified

---

price index. This can be justified because actual inflation matters in deciding about real balances. In the case of interest rates, we did not calculate real interest rates.

[11] Cf. Dickey/Fuller (1979).

[12] Cf. Canova (1998a,b).

[13] Cf. Wolters/Kuhbier/Buscher (1990).

in levels in this analysis were seasonally adjusted using the Berlin Method 4 (BV4). The relevant properties of the indicators are presented in table 2.1.[14]

3. Spectral Analysis

3.1. Methodological Approach[15]

Traditionally, the cyclical properties of time series and lead-lag-structures are determined by cross-correlogram analysis.[16] Overlapping oscillations of different periods and with different amplitudes can distort the properties of the correlogram. High auto-correlation complicates the analysis. Furthermore, new developments in time series analysis show that results depend to a large extent, on the kind of transformations used to attain stationary variables, e.g. trend deviations or growth rates. The regression of independent non-stationary time series leads to spurious regression. At the same time, the specification of non-stationary time series in differences while these series are co-integrated leads to misspecification. The problems become even more complicated as the results are sensitive with regard to the assumed parametric model - a problem widely discussed in conjunction with de-trending procedures.

The first step in our research is the discrimination between series which show a significant relationship with the business cycle in the relevant period and those which do not. Spectral analysis proved to be a helpful tool for this purpose. Analytically, a correlogram can be transformed to the frequency domain using Fourier transformation. The spectra functions indicate the contribution of every frequency component to overall variance. By applying spectral analysis to more than one time series, it is possible to calculate some useful measures such as squared coherence and gain, which allow the underlying relationship between different time series to be assessed. In addition, the coherence is invariant to any kind of linear transformation.

---

[14] All time series specified in levels are seasonally adjusted. Most of them are adjusted using the Berlin method BV4. The Consumer Sentiment Indicator and the Real Effective Exchange Rate were only available on a seasonally adjusted basis (X-11). Regarding the series which had to be specified in growth rates due to their stationarity properties, in most cases we used unadjusted time series where the transformation into annual growth rates served as a seasonal adjustment procedure. The reason for this was an observed bias towards non-stationarity when we transformed already seasonally adjusted time series into growth rates.

[15] Cf. König/Wolters (1972); Wolters/Kuhbier/Buscher (1990); Wolters (1996); Kirchgässner/Wolters (1994); Wolters/Lankes (1989); Koopmans (1974).

[16] Cf. Döpke/Krämer/Langfeldt (1994); Lindlbauer (1995).

Frequency domain analysis allows us to calculate the *squared coherence* as a function of *spectra* and *cross spectra:*

$$K_{xy}(\lambda) = \frac{\left|f_{xy}(\lambda)\right|^2}{f_{xx}(\lambda)f_{yy}(\lambda)}$$

where $f_{aa}(\lambda), a = x, y, u$ is called *spectrum* at frequency λ and $f_{xy}(\lambda)$ is called *cross spectrum* (between x and y).

According to König/Wolters (1972) this is a measure for the stochastic relationship between different components of two processes at specific frequencies.[17]

In other words:

> "...*(T)he squared coefficient of coherence (...) can be interpreted as the proportion of the power at frequency* λ *in either time series (...) which can be explained by its linear regression on the other.*"[18]

Therefore, the measure is comparable to the well-known R^2 in traditional regression analysis. However, the application of frequency domain analysis allows some pitfalls of traditional regressions to be avoided, since coherence is a measure of the degree of *linear association, not of linear dependence*. No causal relationship between the two variables has to be assumed, as is implicitly the case in regression analysis. Furthermore, no specific model needs to be specified for the determination of the direction of dependence. One of the most important advantages is the invariance against any kind of linear transformation, including that to growth rates. It is worth mentioning, that as Kirchgässner and Wolters (1994) have shown, a coherence of one at frequency zero indicates a co-integration relationship between two time series.[19] This finding is in line with a popular interpretation of co-integration in the sense that in the long run (a frequency of zero corresponds to a cycle of infinite length) both time series *are strongly related and do not diverge from each other*. Our coherence estimations can therefore be regarded as an informal test for co-integration relationships between the indicator and the reference series.

Similar to all other non-parametric approaches, the empirical application of spectral analysis has disadvantages as well. Relatively long time series are required to get reliable results. Moreover, the analysis is complicated by the trade-off between bias and variance.

[17] König/Wolters (1972: 120).

[18] Koopmans (1974: 142).

[19] Cf. Kirchgässner/Wolters (1994).

3.2. *Results*

The results of coherence estimation for the business cycle reference series and the indicators are shown in figures 3.1 to 3.5.[20]

The null hypothesis of no significant influence (at a 5 % confidence level) was tested using a significance test statistic developed by Koopmans (1974). The horizontal line in our graphs represents the 5% confidence band.[21] Coherence values *above* this level show *significant association* between these two series at specific frequencies (which were transformed into periods for better understanding, e.g. months).

How can these results be interpreted? In traditional business cycle literature[22], a period of one year up to six or eight years is regarded as relevant for business cycle movements. Hence, coherence tests were carried out for this time period (that is, on the left side of our graphs).

Our tests lead to some interesting conclusions. First, over the whole frequency domain, the indices of net production show a strong association with the reference series (see figure 3.1).[23] The fact that coherence is quite high throughout the frequency spectrum shows that these series are to a large extent identical. This is not surprising, given the overlapping data base. These indices should therefore be used as coincident indicators, therefore it would seem inappropriate to use them as leading indicators. For this reason, we decided to exclude these indicators from our further investigation.

Second, all order inflows as well as the *ifo* climate and expectation indicators show significant coherence in the period under consideration (see figure 3.2). Third, quite interestingly, all monetary aggregates (see figure 3.3 and 3.4) – real and nominal – and the real effective exchange rate are insignificant, whereas both interest rates, as well as their spread[24] have explanatory power, but only with little significance.[25] This is the case for consumer confidence, as well as the spread between government and private bond yields (see figure 3.5).

---

[20] For the empirical estimation we used the program SPEKTRAL, developed at the Freie Universität Berlin, Faculty of Economics, chair of Professor Jürgen Wolters. The following parameters were used: length of the time series: 240 data points, number of estimated function values: 72, covariances: 36. Applying a Parzen window the estimation has 24 degrees of freedom [cf. König/Wolters (1972: 72)]. We thank Professor Jürgen Wolters for sharing the programme files.

[21] Koopmans (1974), annex, table A9.6.

[22] Cf. Zarnowitz (1992).

[23] In the case of manufacturing industry the time series is co-integrated with the reference series. This result can be derived from the coherence between these two series because in this case the coherence is one at zero frequency . Cf. Kirchgässner/Wolters (1994).

[24] In fact, the spread shows only small signs of significance. We decided to include the spread into the further explorations because of its dominant role in leading indicator literature.

[25] In other analyses interest rates show very good indicator properties. Cf. Kirchgässner/Savioz (1998).

These findings correspond with the research of Bernanke and Blinder (1992) or Friedman and Kuttner (1992), who showed that the information content of money - however defined - is to a large extent obsolete if interest rates are taken into account.

In our further research we decided to exclude all monetary aggregates as well as the real effective exchange rate and the production indices and to examine only the relationships between the reference series and indicators with significant coherence in the interval relevant for business cycle research. This strategy was chosen since one of the basic properties which a reliable indicator should possess is that movements in the indicator series should resemble those of the indicator series. In our spectral analysis based test the excluded time series do not fulfil the requirement for the relevant interval.

4. Analysis of Lead-Lag Structures

In the second part of our analysis, we examined the lead-structure between indicator and reference series. Basically, the *phase* can be determined within multivariate frequency domain analysis. However, this procedure has some disadvantages. Owing to the ambiguous nature of trigonometric functions, these measures are difficult to interpret. Furthermore, estimates are imprecise if the coherence is quite small.[26]

As a result we decided to use other techniques. First we used cross-correlograms to identify - via the maximum of the coefficient of correlation calculated at different lags - possible lead-lag-structures. We then asked whether the inclusion of past values of the indicator variable would improve the forecast of the reference series. To achieve this we performed Granger-causality tests.

4.1. Cross Correlation

As a first approximation of lead-lag structures between the chosen reference series and those indicators that passed the spectral analysis criterion, cross correlation can be estimated. Traditionally, the maximum of the coefficient of correlation is seen as the "lead" or "lag" of the indicator in relation to a reference series. But these measures should be interpreted cautiously, as they can be distorted by overlapping oscillations. We estimated the coefficient of correlation between the reference and indicator series with a lag length of 24 on each side. The thin lines in our graphs represent a rough estimation of the 5% significance band.[27] Values outside this band indicate a "significant" correlation. Results are plotted in figures 4.1 to 4.3.

---

[26] Cf. Wolters (1996).

[27] The significance band is calculated as $\pm 2 / \sqrt{T}$, where T is the number of observations.

The indices of order inflows, the *ifo* business climate and the long-term interest rate show strong signs of co-movement. *Ifo* business expectations, the consumer confidence indicator and the spread between the yields of government bonds and private bonds, as well as the interest rate spread show some lead in the period of one to twelve months, which is of special interest for short-term forecasting. The coefficient of correlation for the short-term interest rate shows signs of lagging instead of the expected leading property. To attain more information, we carried out two types of Granger-causality tests in our investigation - a traditional pair-wise Granger-causality test contained in every econometric software package and a modified test.

4.2. Pair-wise Granger-Causality Tests

Determining whether movements in the indicator series "lead" movements in the reference series, is of crucial importance in identifying reliable indicators. Granger-causality tests were developed for the assessment of such questions. The test on Granger-causality attempted to determine whether changes in the indicator series precede changes in the reference series or vice versa: We included past values of a stationary indicator series to a regression of a stationary reference series on its own lagged variables. If the fit improves significantly by this inclusion, the indicator series is Granger-causal.[28]

A common difficulty in performing such tests is the choice of lag length, because the results are not independent from the chosen lag structure.[29] Furthermore, standard econometric software packages carry out these tests with fixed length on both sides, something that is criticised as it may lead to misspecification. We chose a twofold strategy. First we carried out standard pair-wise Granger-causality tests with lags of up to 3, 6 and 12 months on each side. This helped to identify possible "causality" relations in the above-mentioned sense. Then we estimated a univariate equation and added individual lags of the indicator series. We chose the Schwarz information criterion to assess improvements in specification. The second strategy helps to avoid misspecification and serves as a means of determining the lag structure.

The results of the first approach are summarized in table 4.1. The order inflow to producers of investment goods, as well as the order inflow in manufacturing industry and all *ifo* indicators show strong signs of Granger-causality. Short-, as well as long-term interest rates are Granger-

---

[28] "Granger-Causality tests are in fact something of a misnomer: in practice all such tests simply examine whether movements in one variable regularly precede those in another variable. There can be no valid test of true causality on this basis in a world where individuals are forward-looking. A simple example is the purchase of anti-freeze in the months leading up to winter: it is clear that winter causes antifreeze purchases; but a typical Granger-Causality Test would suggest the reverse causation, since the antifreeze purchases come first. However, in the context of the search for potential leading indicators, this problem does not arise: in our example, anti-freeze purchases are a good **leading indicator** of winter."; Salazar et. al. (1996: 50).

[29] Cf. Gujarati (1995: 622).

causal, if a lag structure of three or six months is chosen. Both spreads as well as the consumer confidence are insignificant with a lag length of 6 or 12 months, but slightly significant if a range of three months is chosen.

For most of the relevant indicators, the pair-wise Granger-tests show that causality runs from the indicator to the reference series.

4.3. Individual Granger-Causality Tests

Due to the above-mentioned disadvantages of the standard Granger-causality tests carried out by standard econometric software, we performed individual tests as well.[30] Contrary to the pair-wise tests, we tested for one direction of causality only.

First, we estimated the best univariate specification for the reference series (t-values in parentheses):

$$y = \underset{(1.61)}{0.002} + \underset{(5.43)}{0.33}\,y_{-1} + \underset{(5.05)}{0.31}\,y_{-2} + \underset{(4.17)}{0.27}\,y_{-3} + \underset{(2.04)}{0.12}\,y_{-11} - \underset{(-4.87)}{0.28}\,y_{-12}$$

with:

$$\overline{R}^2 = 0.67$$
$$DW = 1.87$$
$$SIC = -4.51$$

where DW denotes the Durbin-Watson statistic and SIC the Schwartz information criterion. No serial correlation remained in the residuals. In the second step we estimated regressions specified in the form of Granger-tests. Here, we could add individual lags (always one) of the indicator series to the univariate regression. In general, the above-mentioned equation was modified to:

$$y = \beta^0 + \beta^1 y_{-1} + \beta^2 y_{-2} + \beta^3 y_{-3} + \beta^4 y_{-11} + \beta^5 y_{-12} + \gamma^1 x_{-t}$$

where t = 1,2,....24.

The value of the Schwartz information criterion of the latter equation was compared with the value of the Schwartz information criterion of the univariate equation (the dotted line in figures 4.4 to 4.6), for equations from the first up to and including the 24th lag. An improved information criterion in comparison with the information criterion of the univariate estimation was interpreted as a sign of Granger-causality and because only individual lags were used, the absolute minimum of the criterion served as a means of identifying the most significant "lead" between reference and indicator series. The results are summarized in the figures 4.4 to 4.6.

The modified Granger-tests show that the inclusion of the order inflows to producers of investment goods and in manufacturing industry improve the equation in the very short run (see

---

[30] Cf. Wolters (1996).

figure 4.4). In the case of *ifo* indicators, the inclusion of the respective indicator improved the fit in the first month up to half a year (see figure 4.5). The only exceptions were order inflows to and *ifo* business expectations of producers of intermediate input. In these cases, no significant improvement was achieved.

Both interest rates have little additional explanatory power. However, only the inclusion of the interest rate spread improves the fit of equation into the period from 4 to 24 months, but without a clearly defined local minimum (see figure 4.4). The spread between government bond and private bond yields as well as the consumer sentiment indicator improve the fit in the very short run (see figure 4.6).

To a large extent, the results of the individual Granger-tests confirm the results found earlier. It is quite interesting that the inclusion of *ifo* indicators leads to the lowest values of the Schwartz information criterion. This means, since we held all other parameters (number of regressors and estimation period) constant, that compared with all other indicators, their inclusion improved the in-sample forecasting power most.

Because the individual Granger-test did not find causality in the case of order inflows to producers of intermediate inputs, as well as for the interest rates, we decided to exclude these variables from further investigations. But, in the case of *ifo* business expectations of producers of intermediate input, we decided to retain this series, since the pair-wise Granger-test supported the hypothesis of causality.

5. Out-of-Sample Forecasts

One interesting question remains to be answered. Are the indicators with the best in-sample performance also the indicators with the best out-of-sample performance? The answer is by no means obvious.

For most of the indicators examined, the fit of bivariate equations experienced the greatest improvement in the very short run (in most cases, and especially for *ifo* indicators, the most significant lag structure is one or two lags). In this case, exercising out-of-sample forecasts requires forecasts of the exogenous variables, which is sometimes done by AR processes. However, we chose another strategy.

First, we constructed a VAR that includes the reference series and the indicator series. The maximum lag was restricted to 12 months and single VARs were specified according to significant t-values. The specifications of the VARs can be found in table 5.1.

Theil's U is well-known as a measure of forecast accuracy. We calculated a modified Theil's U, as proposed by Davis and Fagan.[31] This measure is defined as the relation of the root mean squared error of a structural forecast (here: the root mean squared error of the VARs or $RMSE^{VAR}$) to the root mean squared error of a "naive" forecast (here: the root mean squared error of the above-mentioned AR-process or $RMSE^{AR}$).

$$\text{Theil's } U = \frac{RMSE^{VAR}}{RMSE^{AR}}$$

A range of Theil's U between zero (perfect prediction) and less than one (a value of one indicates no improvement in comparison with a "naive" forecast) is of special interest for our investigation. Values larger than one can be interpreted as a worsening of the forecast quality compared to the above-mentioned "naive" prognosis. Furthermore, the root mean squared error can be decomposed into a bias, a variance and a covariance proportion.[32] The bias proportion tells us how much the mean of the forecast differs from the mean of the actual series. The variance proportion indicates the differences in variation of the forecast and variation of the actual series. The covariance proportion measures the remaining unsystematic forecasting errors. For a "good" forecast, the bias and variance proportion should be small, whereas most of the remaining errors should concentrate on the covariance proportion.

Regressions were run for the period from 1978:1 to 1990:12. Dynamic three- and six- month forecasts were carried out for the period from 1991 to 1998. In addition, the root mean squared errors for both forecasting methods were calculated and decomposed into bias, variance and covariance. The results are shown in table 5.2. and allow some interesting conclusions to be drawn. Four indicators show satisfactory performance: the order inflow to producers of investment goods, the *ifo* business climate of producers of investment goods and both spreads. The performance of all other indicators was rather dissatisfactory. These results confirm the general scepticism of the usefulness of leading indicator relations.[33]

In the next step we used the four indicators that had performed satisfactorily and included them in a VAR (with a fixed length of three months), which we called "mixed VAR". We then carried out the same forecasts as for the individual indicators. The results were quite satisfactory. Theil's U has an acceptable value and the forecast is unbiased. However, as figure 5.1 shows, the three-month forecast anticipates the business cycle movement better than the six-month forecast, but the inherent inertia of the VARs limits their use in both cases.

---

[31] Davis/Fagan (1997); Döpke (1998).

[32] Cf. Pindyck/Rubinfeld (1998[4]), chapter 8.

[33] As an example for that scepticism concerning Euroland, see Döpke (1998).

6. Conclusion

In our analysis we tested a number of potential indicators using spectral analysis, Granger-tests and out-of-sample forecasts. After each test, we reduced the number of indicators to qualify the rest as reliable leading indicators. The results are satisfactory in that *ifo* indicators, as well as order inflows perform quite well. They show significant coherence in the relevant region, they are qualified by the Granger-tests and in particular the indicators for producers of investment goods, are also qualified by out-of-sample forecasting power. Both spreads show only little significance in the frequency domain, but show signs of Granger-causality and are well qualified by the out-of sample forecast.

The attempt to create an indicator-based VAR (which includes the four best qualified indicators) showed ambiguous results. It works quite well in the very short run of three months, but due to high inertia, this approach has some difficulty in performing six-month forecasts.

To sum up, *ifo* indicators as well as order inflows showed the best results in our tests. Interest rate spreads can also be used as reliable leading indicators. In contrast to other studies monetary aggregates showed a bad performance. Interest rates showed significant coherence in the spectral analysis, but performed badly in the out-of sample performance.

The attempt to create a more sophisticated VAR forecast showed no considerable improvement compared with the bivariate estimations.

In sum, our findings are rather sobering. We found that there are some indicators which improve the forecasts for the very short term significantly. However, in out-of-sample forecasts for every month between 1991 and 1998, which are not included in this article, the values of Theil´s U suggest that no indicator has a stable predictive power. The variance of the forecasting error is in every case very large. Further examinations will have to concentrate on turning points, since publicly used forecasts of annual growth rates depend crucially on the correct prediction thereof.

7. Literatur

Bernanke, B. S.; Blinder, A. S. (1992), The Federal Funds Rate and the Channels of Monetary Transmission. In: American Economic Review, Vol. 83, No. 4, pp. 901-922.

Burns, A. F.; Mitchell, W. C. (1946), Measuring Business Cycles, New York.

Canova, F. (1998a), Detrending and Business Cycle Facts. In: Journal of Monetary Economics, Vol. 41, pp. 475-512.

Canova, F. (1998b), Detrending and Business Cycles Facts: A User's Guide. In: Journal of Monetary Economics, Vol. 41, pp. 522-540.

Davis, E. P.; Fagan, G. (1997), Are Financial Spreads useful Indicators of Future Inflation and Output Growth in EU Countries? In: Journal of Applied Econometrics, Vol. 12, No. 6, pp. 701-714.

De Leeuw, F. (1991), Toward a theory of leading indicators, In: Lahiri, K.; Moore, G. H. (eds.), Leading economic indicators: New Approaches and Forecasting Records, pp. 15-56.

Dickey, D.A.; Fuller, W.A. (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, In: Journal of the American Statistical Association, Vol. 74, pp. 427-431.

Döpke, J. (1997), Die Vorratsveränderungen im Rahmen von Konjunkturprognosen, In: Konjunkturpolitik, Vol. 43, No. 4, pp. 344-370.

Döpke, J. (1998), Leading Indicators for Euroland's Business Cycle. Kiel.

Döpke, J.; Krämer, J. W.; Langfeldt, E. (1994), Konjunkturelle Frühindikatoren in Deutschland. In: Konjunkturpolitik, Vol. 40 (1994), No. 2, pp. 135-153.

Entorf, H. (1990), Multisektorale Konjunkturanalyse: Theorie, Empirie und Frühindikatoren "Realer Konjunkturzyklen", Frankfurt and New York.

Friedman, B.; Kuttner, K. (1992), Money, Prices, and Interest Rates. In: American Economic Review, Vol. 83, No. 3, pp. 472-492.

Funke, N. (1997), Predicting Recessions: Some Evidence for Germany. In: Weltwirtschaftliches Archiv, Vol. 133, pp. 90-102.

Gujarati, D. (1995), Basic Econometrics, 3rd edition, New York.

Haberler, G. (1948), Prosperität und Depression, 2nd edition, Bern.

Hayek, F. A. (1931), Prices and Production, London.

Kirchgässner, G.; Savioz, M. (1998), Monetary Policy and Forecasts for Real GDP Growth: An Empirical Investigation for the Federal Republic of Germany. University of St. Gallen (= Discussion Paper 9813).

Kirchgässner, G.; Wolters, J. (1994), Frequency Domain Analysis of Euromarket Interest Rates. In: Kaehler, J.; Kugler, P. (eds.), Econometric Analysis of Financial Markets, Heidelberg, pp. 89-103.

Klein, Ph. A. (1997), The Theoretical Basis of Cyclical Indicators, In: Oppenländer, K. H (ed.) (1997), op. cit., pp. 35-48.

Köhler, K. (1994), Die Spannungszahl – Ein Indikator der gesamtwirtschaftlichen Entwicklung. Berlin.

König, H.; Wolters, J. (1972), Einführung in die Spektralanalyse ökonomischer Zeitreihen. Meisenheim am Glan.

Koopmans, L. H. (1974), The Spectral Analysis of Time Series. New York.

Koopmans, T. (1947), Measurement without Theory, In: Review of Economics and Statistics, Vol. 29, pp. 161-172.

Langfeldt, E. (1994), Die Zinsstruktur als Frühindikator für Konjunktur und Preisentwicklung in Deutschland. Kiel (= Kieler Arbeitspapier. 615).

Langmantel, E. (1999), Das ifo Geschäftsklima als Indikator für die Prognose des Bruttoinlandsprodukts. In: ifo Schnelldienst, Vol. 52, No. 16-17, pp. 16-21.

Lindlbauer, J.D. (1995), Beurteilung ausgewählter Frühindikatoren – Theorie und Evidenz, In: Oppenländer, Karl Heinrich (ed.), Frühindikatoren. München, pp. 144-176.

Mishkin, F. (1981), The Real Interest Rate: An Empirical Investigation, In: Carnegie-Rochester Series on Public Policy. Vol. 15, pp. 151-200.

Moore, G.H.; Zarnowitz, V. (1986), The Development and Role of the National Bureau of Economic Research's Business Cycle Chronologies. In: Gordon, R. J. (ed.), The American Business Cycle. Continuity and Change. Chicago, pp. 735-779.

Oppenländer, K. H. (ed.) (1997), Business Cycle Indicators. Aldershot.

Pindyck, R. S.; Rubinfeld D. L. (1998), Econometric Models and Economic Forecasts, 4th edition, New York.

Salazar, Eduardo et. al. (1996), Leading Indicators. Paris: OECD.

Sauer, Ch.; Scheide, J. (1995), Money, Interest Rate Spreads, and Economic Activity. In: Weltwirtschaftliches Archiv, Vol. 131, pp. 708-722.

Seifert, M. (1999), Frühindikatoren der Konjunktur für die Europäische Währungsunion. Halle (= Diskussionspapier des Instituts für Wirtschaftsforschung. 87).

Stock, J.; Watson, M. (1989), New Indexes of Coincident and Leading Economic Indicators. In: Blanchard, O.; Fischer, St. (eds.), NBER Macroeconomics Annual, Cambridge Mass., pp. 351-394.

Tichy, G. (1994), Konjunktur: Stilisierte Fakten, Theorie, Prognose. Berlin et. al.

Wagemann, E. (1928), Konjunkturlehre, Berlin.

Wolters, J. (1996), Zur Beurteilung verschiedener Frühindikatoren für die Produktionsentwicklung, In: Helmstädter, E.; Poser, G.; Ramser, H. J. (1996), Beiträge zur angewandten Wirtschaftsforschung: Festschrift für Karl Heinrich Oppenländer. Berlin, pp. 339-361.

Wolters, J.; Kuhbier, P.; Buscher; H. S. (1990), Die konjunkturelle Entwicklung in der Bundesrepublik Deutschland: Ergebnisse von Schätzungen alternativer aggregierter Konjunkturmodelle. Frankfurt.

Wolters, J.; Lankes, F. (1989), Das Ifo-Geschäftsklima als Konjunkturindikator. Empirische Ergebnisse für die Verbrauchs- und Investitionsgüterindustrie 1975-1986 In: IFO-Studien, 35. Jg., pp. 198-209.

Zarnowitz. V. (1992), Business Cycles: Theory, History, Indicators, and Forecasting. Chicago.

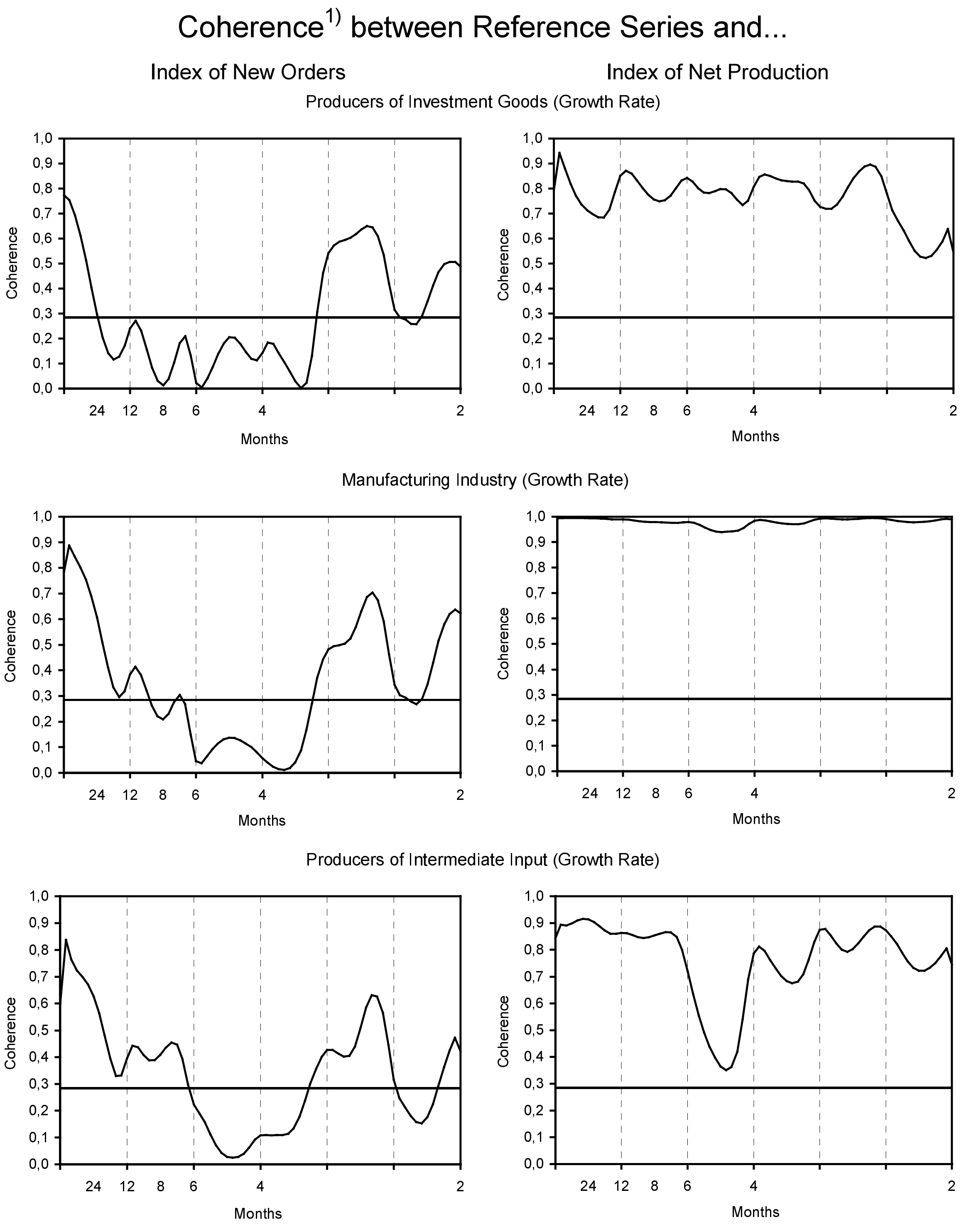
Appendix

Table 2.1: Indicator Properties

Indicator	Integration	Transformation	Source
Index of New Order			
Producers of Investment Goods	I(1)	Annual growth rates[1)]	Eurostat
Manufacturing Industry	I(1)	Annual growth rates	Eurostat
Producers of Intermediate Input	I(1)	Annual growth rates	Eurostat
Index of Net Production			
Producers of Investment Goods	I(1)	Annual growth rates	Eurostat
Manufacturing Industry	I(1)	Annual growth rates	Eurostat
Producers of Intermediate Input	I(1)	Annual growth rates	Eurostat
Ifo Business Expectations			
Producers of Investment Goods	I(0)	Level	Ifo Institute Munich
Manufacturing Industry	I(0)	Level	Ifo Institute Munich
Producers of Intermediate Input	I(0)	Level	Ifo Institute Munich
Ifo Business Climate			
Producers of Investment Goods	I(0)	Level	Ifo Institute Munich
Manufacturing Industry	I(0)	Level	Ifo Institute Munich
Producers of Intermediate Input	I(0)	Level	Ifo Institute Munich
Nominal Money Supply			
M1	I(1)	Annual growth rates	Bundesbank
M2	I(1)	Annual growth rates	Bundesbank
M3	I(1)	Annual growth rates	Bundesbank
M3 enlarged	I(1)	Annual growth rates	Bundesbank
Real Money Supply			
M1	I(1)	Annual growth rates	Bundesbank
M2	I(1)	Annual growth rates	Bundesbank
M3	I(1)	Annual growth rates	Bundesbank
M3 enlarged	I(1)	Annual growth rates	Bundesbank
Real Credit Supply[2)]	I(1)	Annual growth rates	Bundesbank
Short-Term Interest Rate (3 month FIBOR)	I(1)	Annual growth rates	Bundesbank
Long-Term Interest Rate (*Umlaufsrendite*)	I(1)	Annual growth rates	Bundesbank
Interest Rate Spread	I(0)	Level	Bundesbank
Real Effective Exchange Rate	I(0)	Level	OECD
Spread between Government and Private Bond Yields	I(0)	Level	Bundesbank
Consumer Sentiment Indicator	I(0)	Level	OECD

1) Annual growth rates = log(x)-log(x(-12)).- 2) Nominal credit supply was excluded from further analysis because the annual growth rate remained I(1).

Figure 3.1



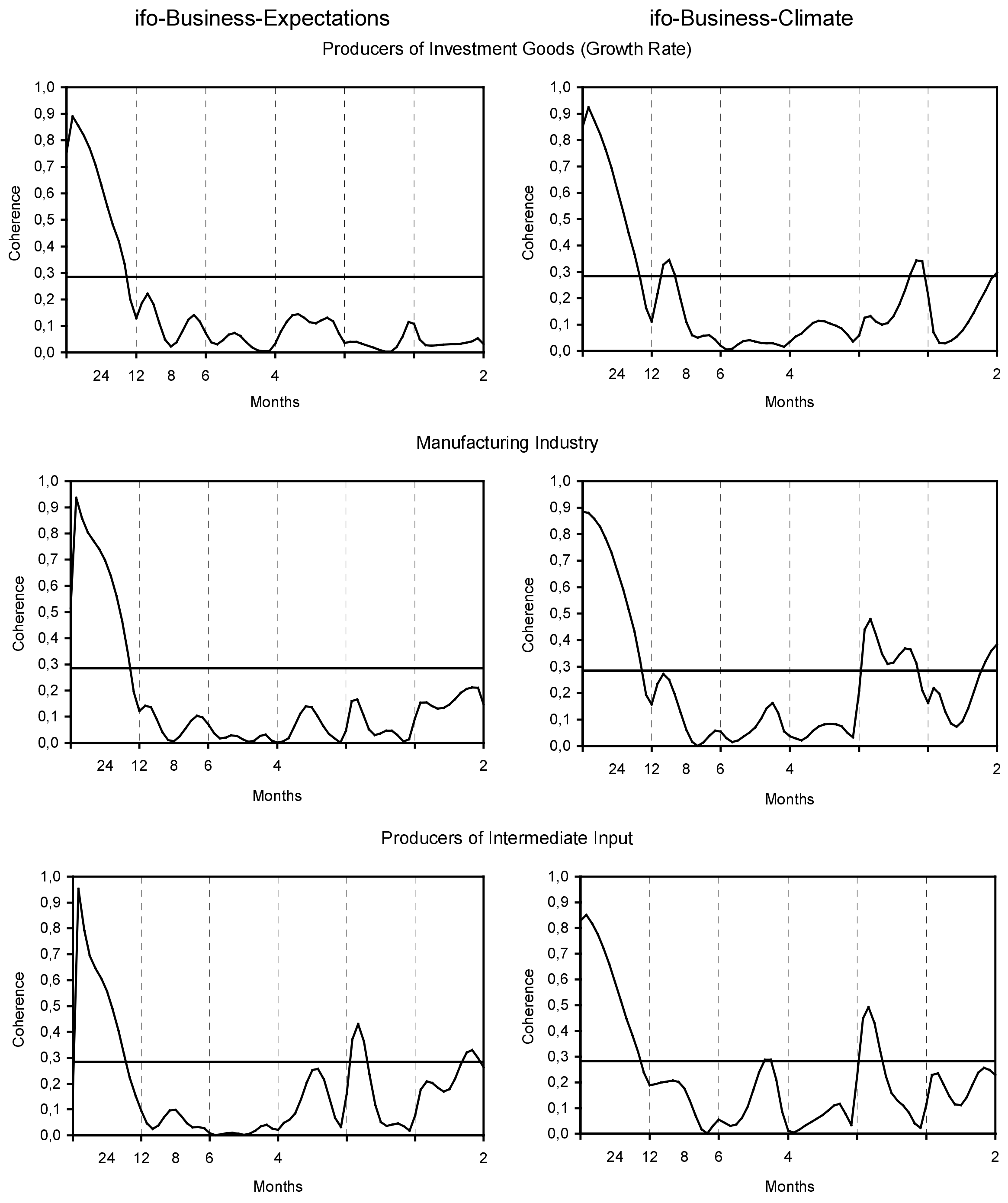
1) The coherence was estimated using 72 datapoints and 36 covariances which implies 24 degrees of freedom (Parzen window).

Source: Calculations of DIW.

Deutsches Institut für Wirtschaftsforschung

Figure 3.2

Coherence[1)] between Reference Series and...



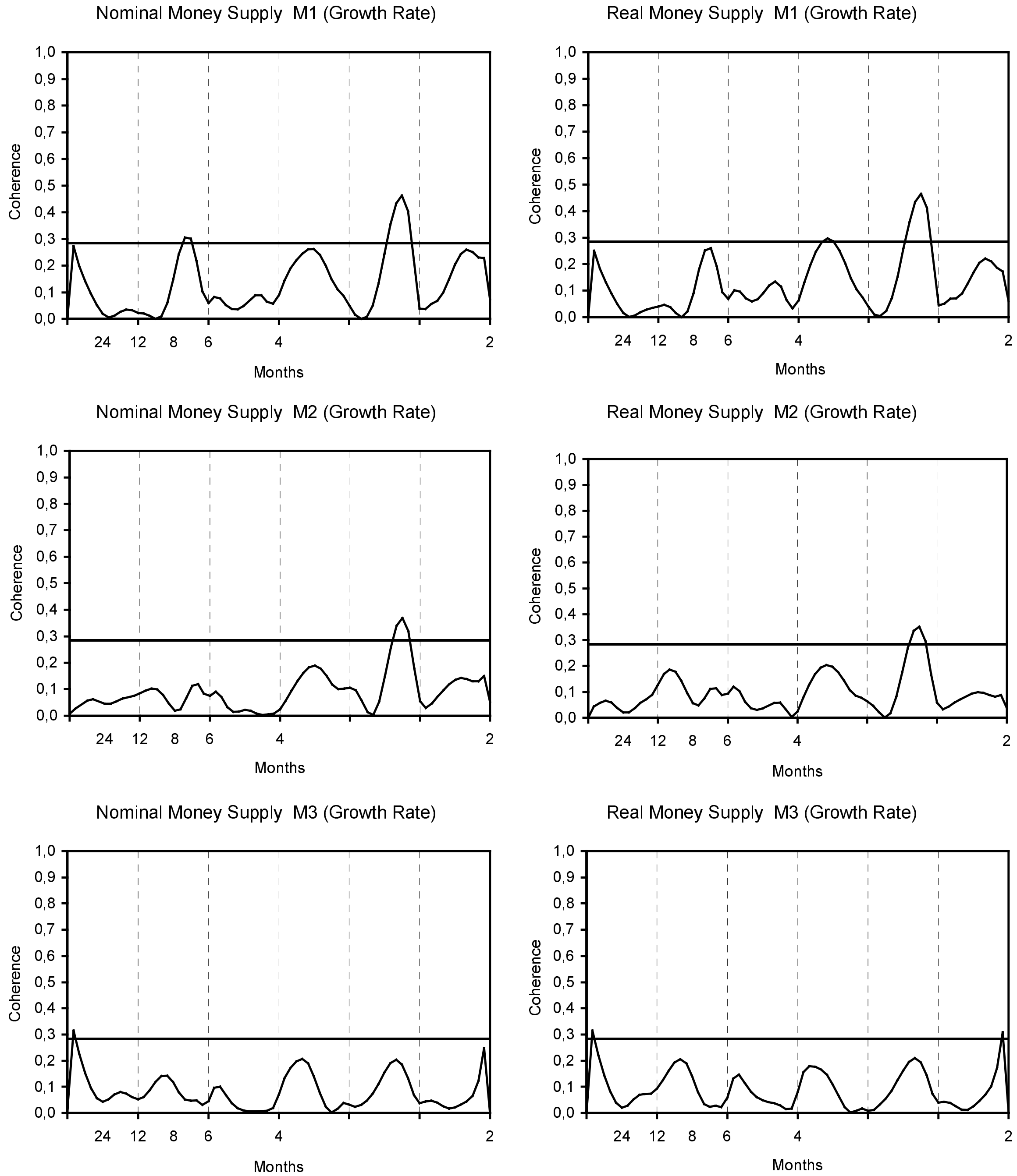
1) The coherence was estimated using 72 datapoints and 36 covariances which implies 24 degrees of freedom (Parzen window).

Source: Calculations of DIW.

DIW Deutsches Institut für Wirtschaftsforschung

Figure 3.3

Coherence[1)] between Reference Series and...

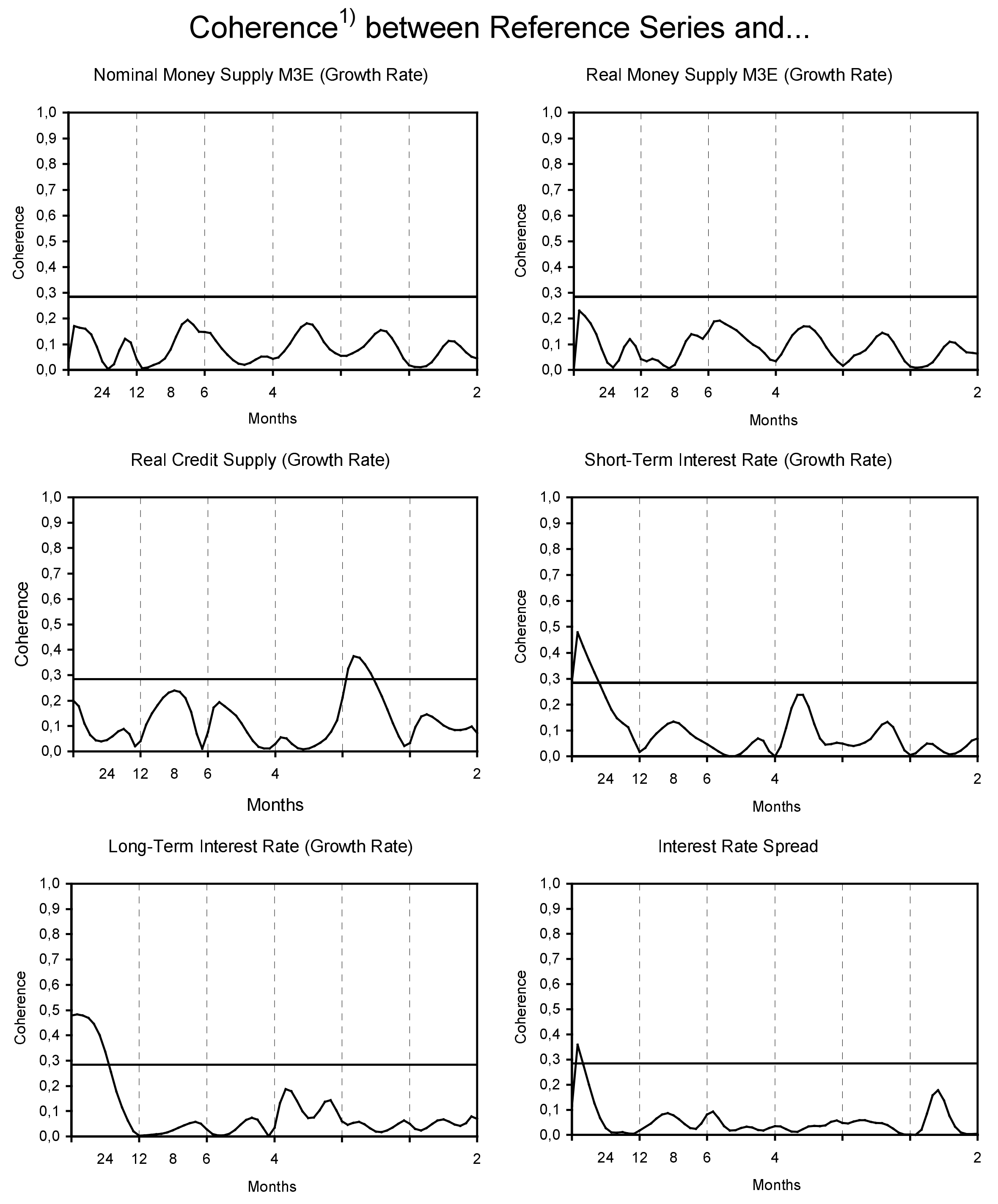


1) The coherence was estimated using 72 datapoints and 36 covariances which implies 24 degrees of freedom (Parzen window).

Source: Calculations of DIW.

DIW Deutsches Institut für Wirtschaftsforschung

Figure 3.4

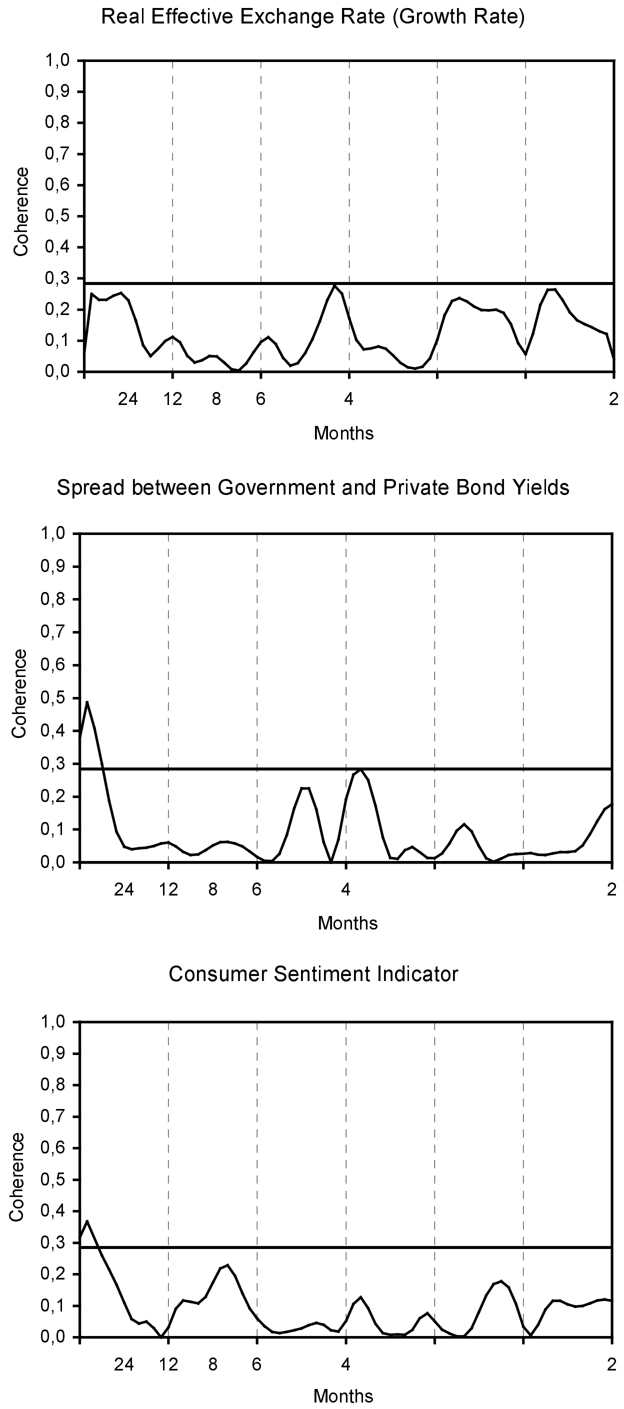


1) The coherence was estimated using 72 datapoints and 36 covariances which implies 24 degrees of freedom (Parzen window).

Source: Calculations of DIW.

Figure 3.5

Coherence[1] between Reference Series and...



1) The coherence was estimated using 72 datapoints and 36 covariances (24 degrees of freedom).

Source: Calculations of DIW.

Deutsches Institut für Wirtschaftsforschung

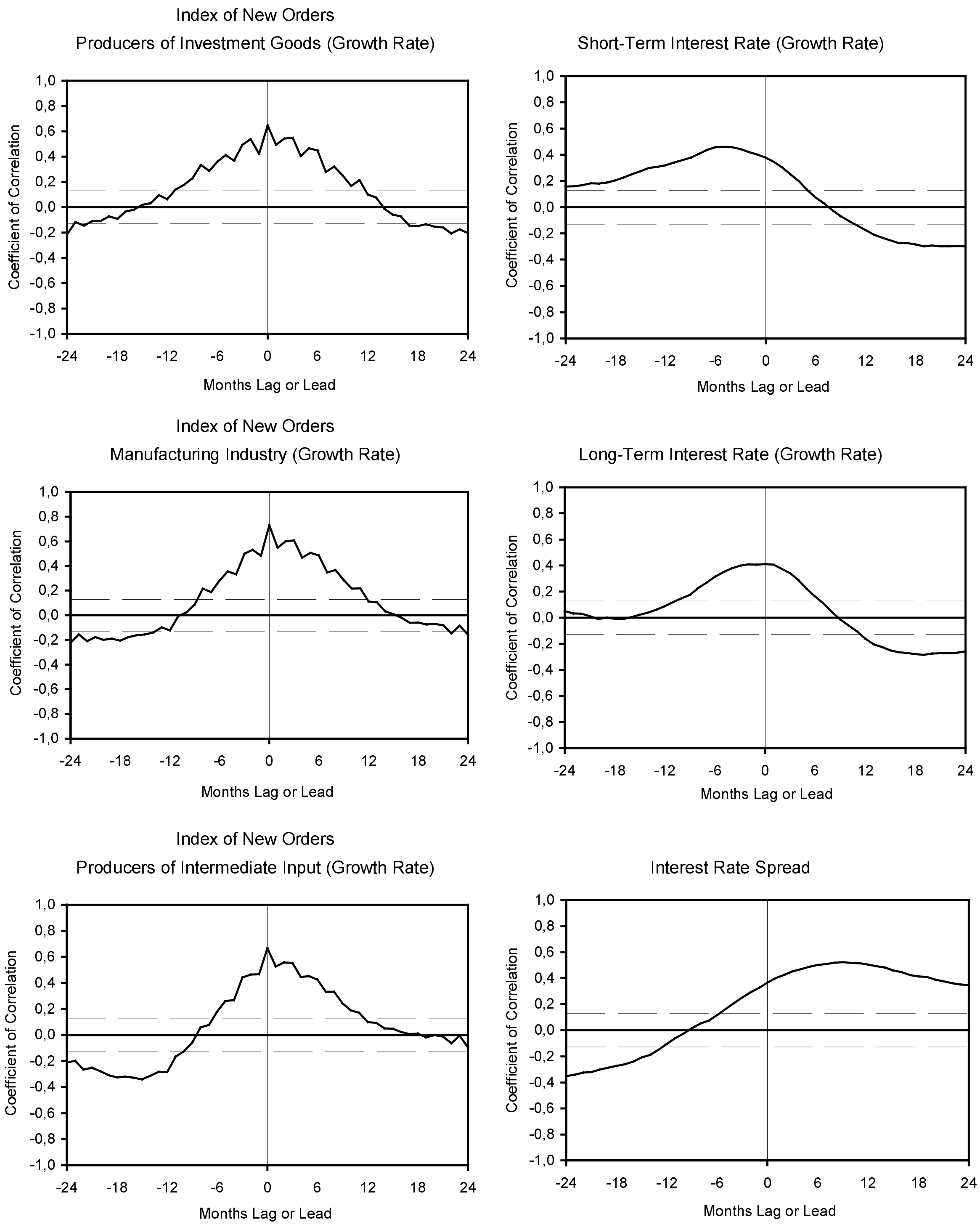
Table 4.1: Pairwise Granger-Causality Tests

Y (reference series)	X (indicator series)	F (X→Y)			F (Y→X)		
		3 Months	6 Months	12 Months	3 Months	6 Months	12 Months
Industrial Production without Construction (growth rate)	Index of New Orders, Producers of Investment Goods	3.40**	2.89***	2.55***	7.13***	4.84***	2.97***
	Index of New Orders, Manufacturing Industry	4.38***	2.89***	1.98**	0.85	2.23**	1.89**
	Index of New Orders, Producers of Intermediate Input	1.90	1.82*	1.47	0.48	1.08	1.77*
	Ifo Business Climate, Producers of Investment Goods	18.75***	10.80***	5.65***	2.22*	1.77	1.18
	Ifo Business Climate, Manufacturing Industry	15.92***	8.37***	4.88***	2.22*	2.71**	1.51
	Ifo Business Climate, Producers of Intermediate Input	10.91***	5.40***	4.18***	0.62	2.70**	1.83**
	Ifo Business Expectations, Producers of Investment Goods	17.88***	9.65***	2.72***	1.87	0.49	1.39
	Ifo Business Expectations, Manufacturing Industry	12.54***	7.53***	4.03***	3.17**	1.76	0.68
	Ifo Business Expectations, Producers of Intermediate Input	5.58***	4.40***	2.72***	4.52**	3.36**	1.39
	Short-term Interest Rate	2.69**	2.58**	1.08	2.14*	0.98	1.24
	Long-term Interest Rate	2.24*	2.20**	1.09	1.29	1.20	0.77
	Interest Rate Spread	2.93**	1.51	0.92	1.45	1.14	1.15
	Spread between Government and Private Bond Yields	3.03**	1.58	1.65*	0.15	0.18	0.91
	Consumer Sentiment Indicator	2.46*	1.60	1.20	0.16	0.99	0.70

Note: ***,** and * denote significance at the 1, 5 and 10 per cent level.

Figure 4.1

Cross Correlations between Reference Series and ...



Source: Calculations of DIW.

DIW Deutsches Institut für Wirtschaftsforschung

Figure 4.2

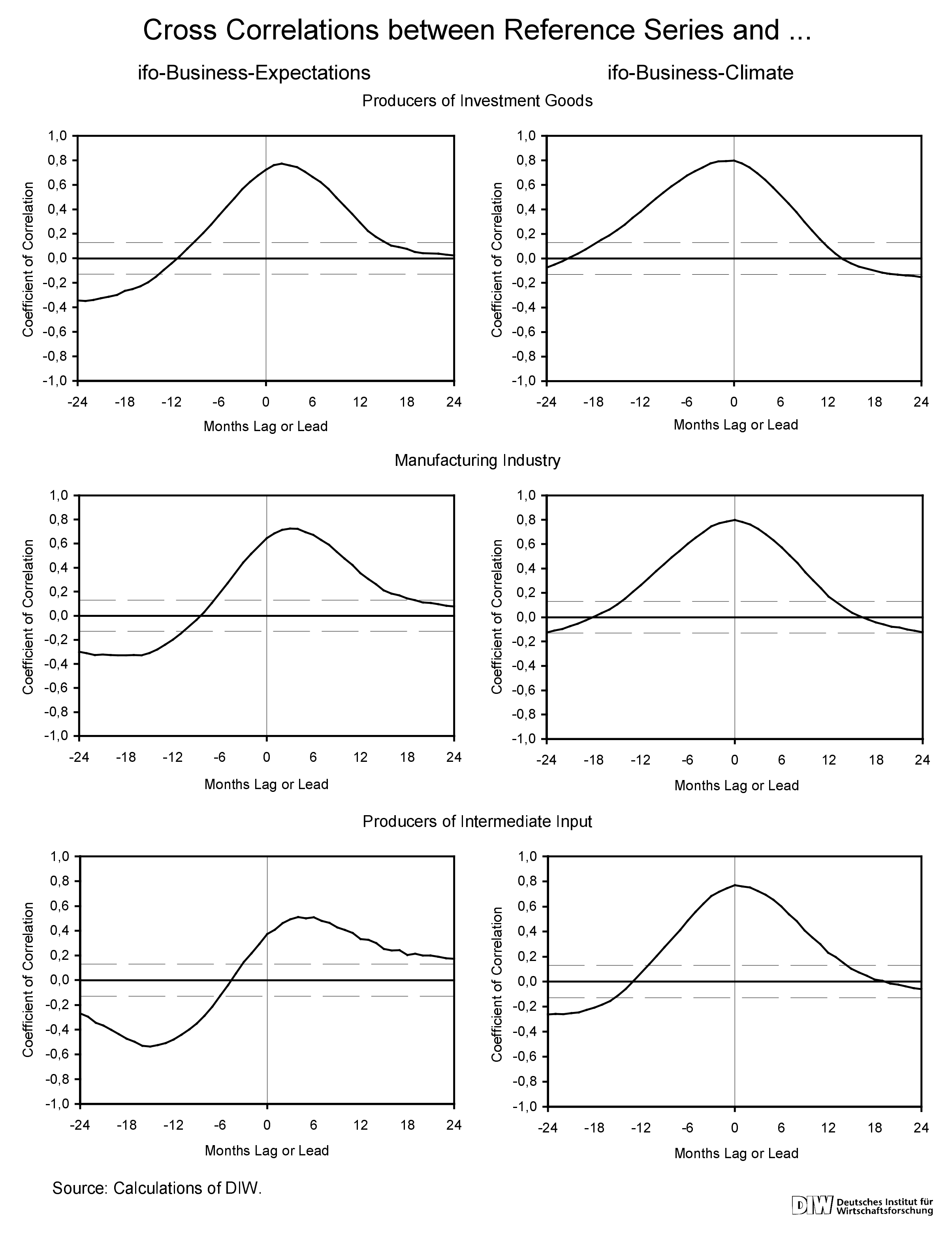
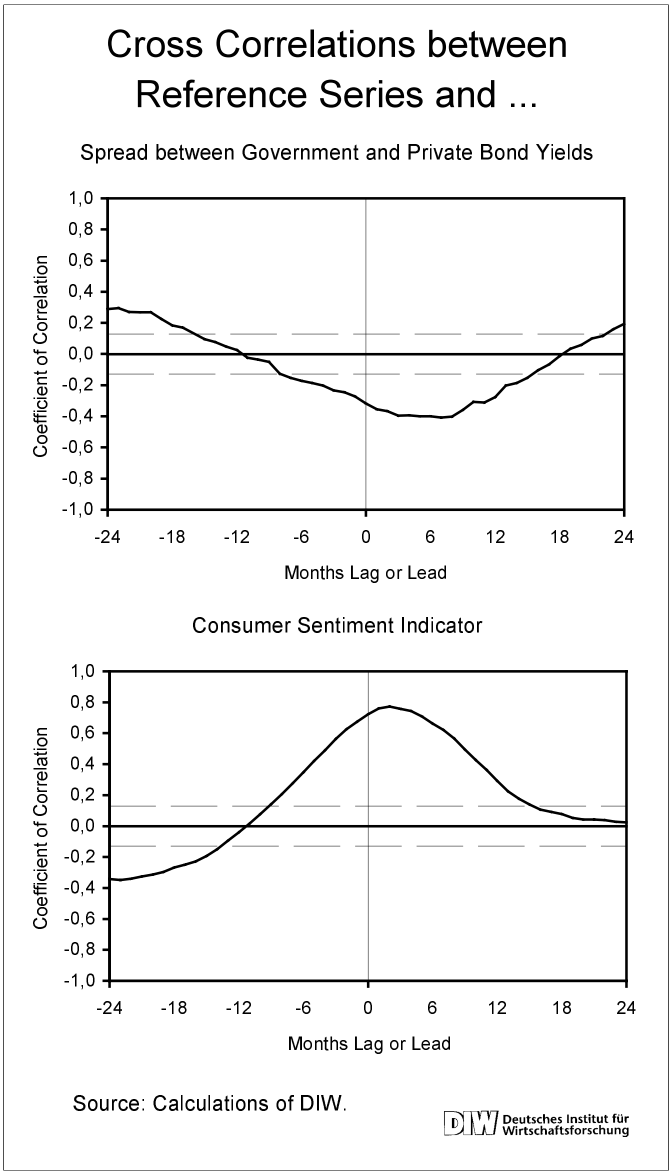

Cross Correlations between Reference Series and ...
ifo-Business-Expectations
ifo-Business-Climate
Producers of Investment Goods
Manufacturing Industry
Producers of Intermediate Input
Coefficient of Correlation
Months Lag or Lead
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
-24
-18
-12
-6
0
6
12
18
24
Source: Calculations of DIW.
DIW Deutsches Institut für Wirtschaftsforschung

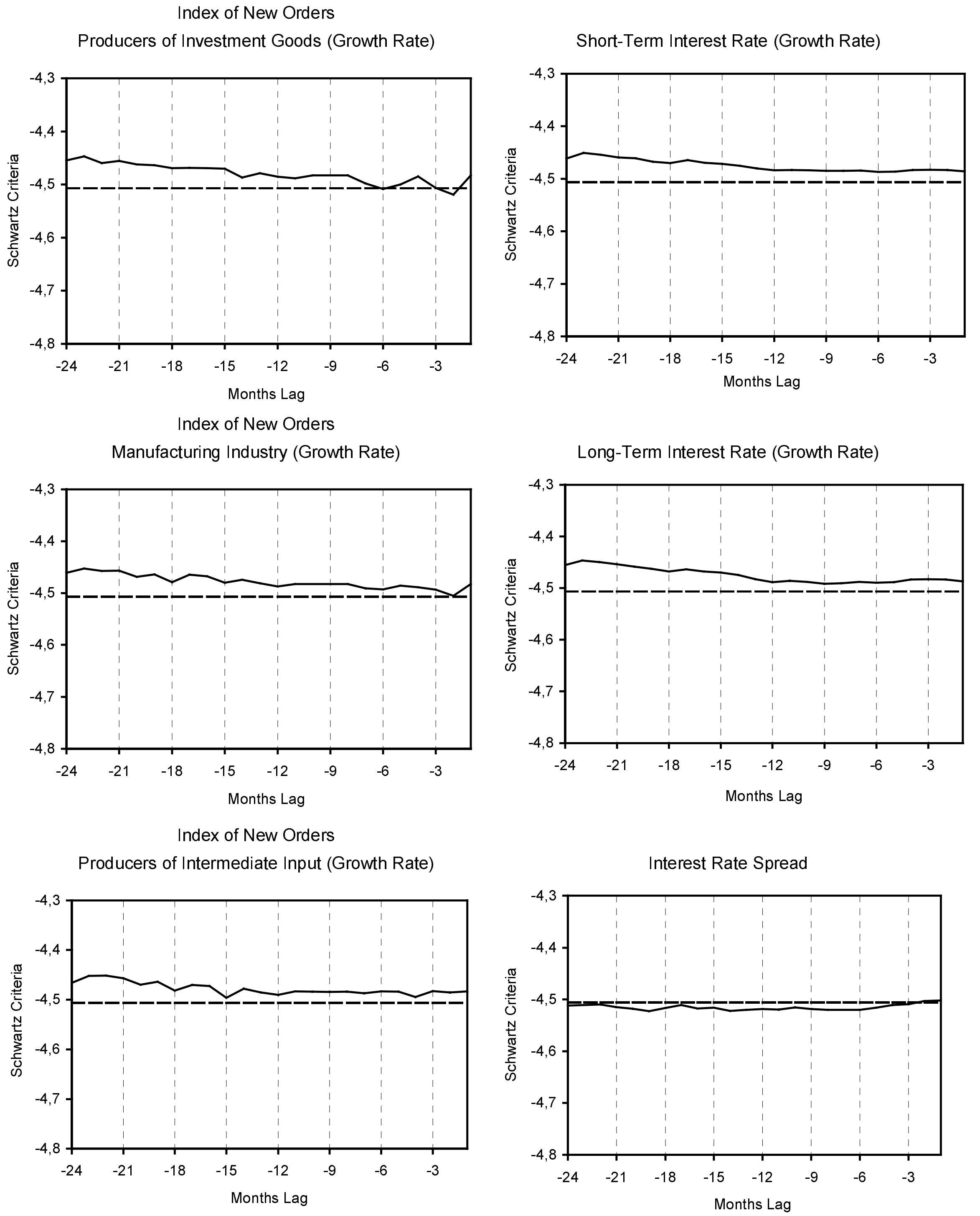
Figure 4.3



Cross Correlations between
Reference Series and ...
Spread between Government and Private Bond Yields
Coefficient of Correlation
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
-24 -18 -12 -6 0 6 12 18 24
Months Lag or Lead
Consumer Sentiment Indicator
Coefficient of Correlation
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
-24 -18 -12 -6 0 6 12 18 24
Months Lag or Lead
Source: Calculations of DIW.
Deutsches Institut für Wirtschaftsforschung

Figure 4.4

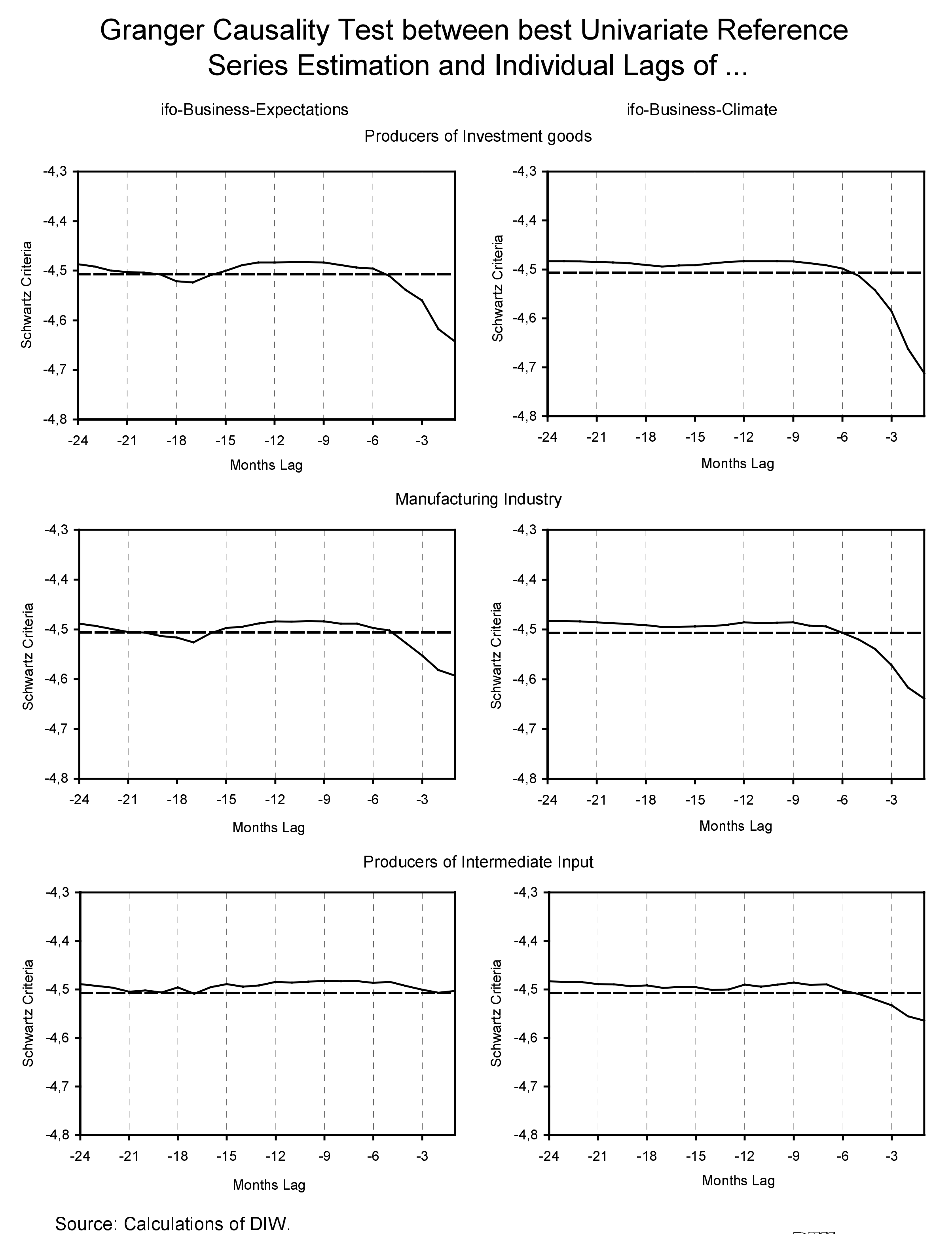
Granger Causality Test between best Univariate Reference Series Estimation and Individual Lags of ...



Source: Calculations of DIW.

Deutsches Institut für Wirtschaftsforschung

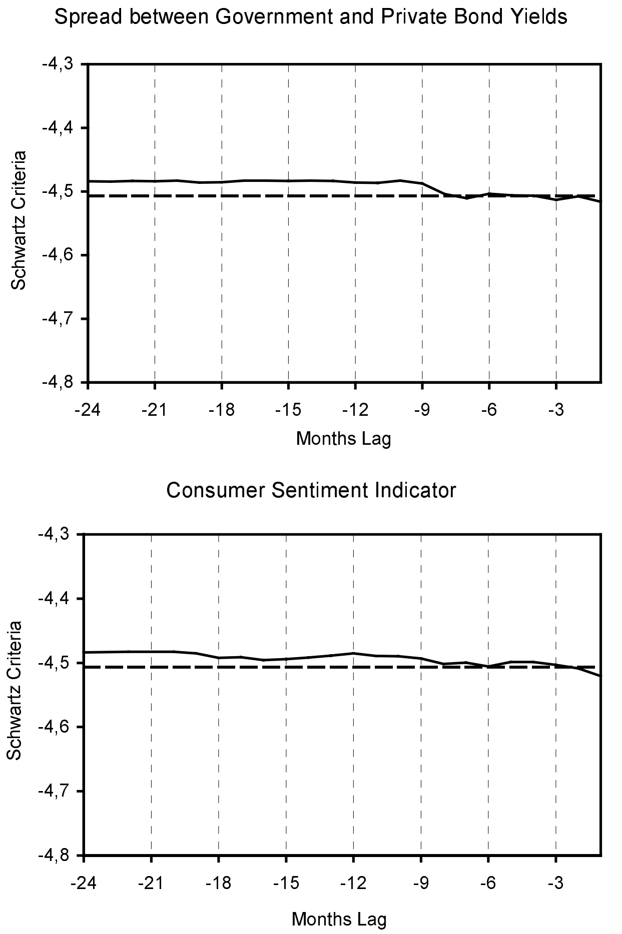
Figure 4.5



Source: Calculations of DIW.

Figure 4.6

Granger Causality Test
between best Univariate
Reference Series Estimation
and Individual Lags of ...



Source: Calculations of DIW.

Deutsches Institut für Wirtschaftsforschung

Table 5.1: Specification of VARs

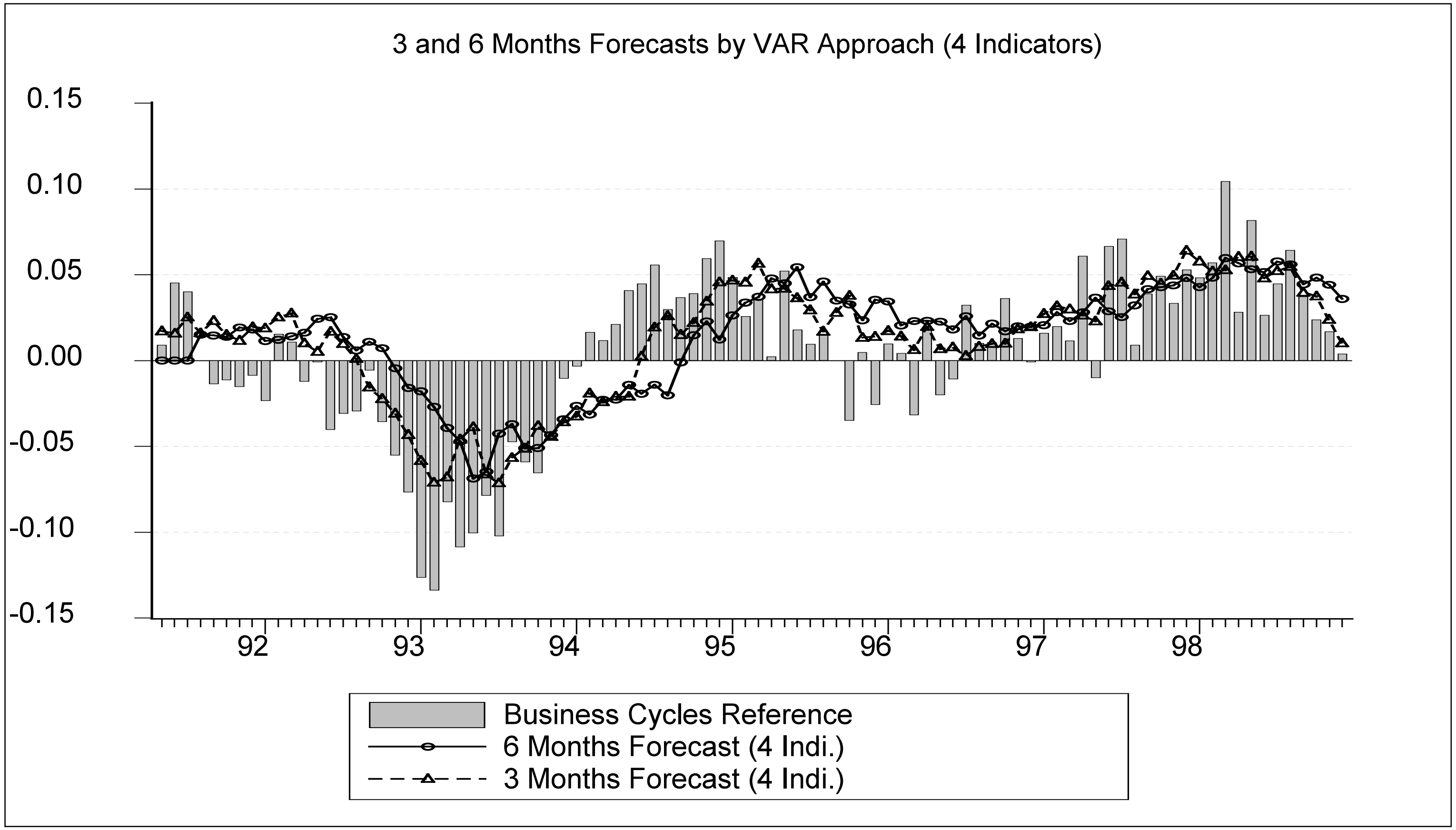
Estimated (1978:1 to 1990:12) VAR between the Reference Series and...	**Lag Specification**
Index of New Orders, Producers of Investment Goods	1-3, 6-7, 12
Index of New Orders, Manufacturing Industry	1, 3, 12
Ifo Business Climate, Producers of Investment Goods	1, 3, 5, 12
Ifo Business Climate, Manufacturing Industry	1, 3, 5, 12
Ifo Business Climate, Producers of Intermediate Input	1-3, 5-6, 12
Ifo Business Expectations, Producers of Investment Goods	1-3, 12
Ifo Business Expectations, Manufacturing Industry	1, 3, 5, 12
Ifo Business Expectations, Producers of Intermediate Input	1-3, 5, 7, 12
Interest Rate Spread	1-3, 12
Spread between Government and Private Bond Yields	1-3, 8-12
Consumer Sentiment Indicator	1-3, 12

Table 5.2: Out-of-sample Forecast Results for 1991-1998

VAR between Business Cycle Reference Series and...	3 Months VAR Forecast				6 Months VAR Forecast			
	Modified Theil's U 1)	Forecast Measures of RMSE ^{VAR}			Modified Theil's U 1)	Forecast Measures of RMSE ^{VAR}		
		Bias Proportion	Variance Proportion	Covariance Proprtion		Bias Proportion	Variance Proportion	Covariance Proportion
Order Inflow, Producers of Investment Goods	0.92	0.051	0.15	0.80	0.94	0.090	0.18	0.73
Order Inflow, Manufacturing Industry	1.02	0.065	0.33	0.61	0.99	0.098	0.41	0.49
Ifo Business Climate, Producers of Investment Goods	0.79	0.13	0.17	0.69	0.80	0.12	0.18	0.70
Ifo Business Climate, Manufacturing Industry	0.99	0.24	0.29	0.47	0.94	0.21	0.25	0.54
Ifo Business Climate, Producers of Intermediate Inputs	1.39	0.50	0.15	0.35	1.31	0.48	0.11	0.41
Ifo Business Expectations, Producers of Investment Goods	1.02	0.43	0.013	0.56	1.00	0.39	0.018	0.59
Ifo Business Expectations, Manufacturing Industry	0.97	0.37	0.13	0.50	0.99	0.42	0.15	0.43
Ifo Business Expectations, Producers of Intermediate Inputs	0.99	0.18	0.12	0.70	1.01	0.24	0.19	0.58
Interest Rate Spread	0.91	0.099	0.27	0.63	0.86	0.16	0.35	0.49
Spread between Government Bonds and Private Bonds	0.94	0.064	0.16	0.77	0.88	0.12	0.27	0.61
Consumer Condidence Indicator	1.00	0.025	0.26	0.74	0.98	0.00001	0.33	0.67
"Mixed VAR" (4 Indicators)	0.86	0.073	0.19	0.74	0.90	0.08	0.18	0.73

1) Modified Theil's U is defined as RMSE ^{VAR} / RMSE ^{AR}.

Figure 5.1


3 and 6 Months Forecasts by VAR Approach (4 Indicators)
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
92
93
94
95
96
97
98
Business Cycles Reference
6 Months Forecast (4 Indi.)
3 Months Forecast (4 Indi.)