DIW Berlin

German Institute
for Economic Research

Discussion Papers

498

Ulrich Fritsche*
Jörg Döpke**

Forecast errors and the macroeconomy – a non-linear relationship?

Berlin, Juni 2005

* DIW Berlin, ufritsche@diw.de

** Deutsche Bundesbank, Economics Department, joerg.doepke@bundesbank.de

IMPRESSUM

DIW Berlin
Deutsches Institut für Wirtschaftsforschung
Königin-Luise-Str. 5
14195 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

DIW Berlin

Discussion Papers 498

Ulrich Fritsche*

Jörg Döpke**

Forecast errors and the macroeconomy – a non-linear relationship?

Abstract: The paper analyses the reasons for departures from strong rationality of German business cycle forecasts based on annual observations from 1963 to 2004. We rely on forecasts from the joint forecast of the so-called "six leading" forecasting institutions in Germany. We test for a non-linear relation between forecast errors and macroeconomic fundamentals and find evidence for such a non-linearity for inflation forecasts. Evidence from probit models further suggests that some macroeconomic fundamentals – especially monetary factors – correlate to large positive or negative forecast growth and inflation forecast errors.

JEL-classification: E32, E37, C52, C53

Keywords: forecast error evaluation, non-linearities, business cycles

Berlin, April 2005

* Corresponding author: Ulrich Fritsche, German Institute of Economic Research, Königin-Luise-Strasse 5, D-14195 Berlin, Germany. Tel +49 30 89789 315; Fax +49 30.89789.102; E-mail: ufritsche@diw.de.

** Deutsche Bundesbank, Economics Department, Wilhelm-Epstein-Strasse 14. Frankfurt. Tel +49-69-9566-3051; Fax +49-69-9566-4317; E-mail: joerg.doepke@bundesbank.de. We would like to thank Jürgen Wolters, Boriss Siliverstovs, Jirka Slacalek, Jonas Dovern, the participants of a seminar at the Free University Berlin as well as participants of the 5th Workshop on Macroeconometrics in Halle/Saale for helpful comments. The views presented in this paper are those of the authors and do not necessarily reflect the views of the Deutsche Bundesbank or DIW Berlin.

Content

Content of Tables

1 Introduction

An overwhelmingly broad area of research reflects on the amount of forecast errors made by professional forecasters (see, for example, Fildes and Stekler, 2002, for a recent survey). To a large degree such studies report departures from the idea of strong rationality of these forecasts. In particular, while the forecasts are mostly found to be unbiased, most recent papers find them to be inefficient (see Döpke and Fritsche, 2004, Hagen and Kirchgässner, 2001, for German data). These findings have led some researchers to the conclusion that forecasters may have other goals than simply making the smallest possible forecast error (see, for example, Laster, Bennet and Geoum, 1999). Another possible explanation may be seen in an asymmetric loss function of the forecasters (Elliot, Komunjer and Timmermann, 2004). In this paper, we offer yet another possible explanation for the reported inefficiency of forecasts.

We argue that the standard result of at least inefficient forecasts may rest on some influential outliers. In other words: seldom, but large forecast errors might be at the root of the alleged irrationality of macroeconomic forecasts. This idea is also consistent with some prominent stylised facts of forecast errors. For example, Heilemann (1998) as well as Loungani (2000), among others, document that forecasters have usually missed upcoming recessions. These findings point to the possibility that large forecast errors occur as rare events, which are seen as a major challenge for forecasters (see, for example, Hendry and Clements, 2001). Furthermore, we argue that this may be caused by the fact that the relationship between macro variables – especially monetary variables – that are usually used to base a forecast and the forecast errors is non-linear. An explanation for this might be found in the literature on how expectations are formed (Ball and Croushore, 2003).

The first step in our empirical research is, therefore designed to identify outliers in equations usually used to test for rationality. In a second step, we test for non-linearities in the relation between some macroeconomic fundamentals and forecast errors. The estimations refer to the forecasts of so-called "six leading" economic research institutions in Germany (the "joint analysis") made in autumn of each year for the subsequent calendar year. These forecasts are probably the most recognised predictions in Germany and, furthermore, are in themselves a sort of a consensus forecast. (see Stäglin, 1998, for a description of the "joint forecast").[1]

Our results indicate that reducing the influence of "outliers" by robust estimation methods still leads to the rejection of the rationality hypothesis. Turning to a possible non-linearity, we use RESET tests, the arranged regression test of Tsay (1989), and the threshold test as described in Hansen (1997). Our main finding is that non-linearities may indeed help to understand the errors made by the institutes. For example, we find a non-linear relation between lagged economic variables, such as monetary ones, and the occurrence of a forecast error. To check whether a rather simple form of non-linearity is appropriate or not, we also check whether the relation differs between overestimations and underestimations. This kind of analysis relies on probit models. From our findings, we conclude that incorporating non-linearities into the forecast models is one of the prime candidates in improving the forecast record.

The paper is organised as follows. Section 2 reviews the evidence on the sources of forecast errors made by professional forecasters with a special emphasis on German data. Section 3 analyses the real-time forecast errors of the six institutes and presents the results of standard tests for rationality of the forecasts. Section 4 presents the results of the non-linearity tests. Section 5 shows the findings based on probit models. A final section provides conclusions.

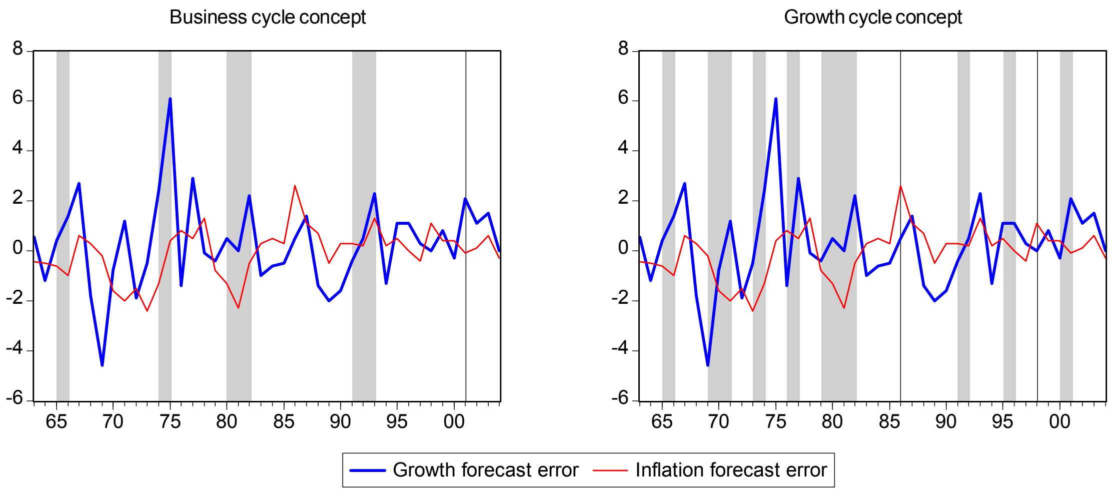
2 Sources of forecast errors

Most studies on business cycle forecasts in Germany conclude that, in general, the forecasts are unbiased, but not necessarily efficient (see Döpke and Fritsche, 2004, for a survey on recent evidence regarding to German data). Döpke and Langfeldt (1995) present evidence supporting the idea that output is generally underestimated in an upswing and overestimated during a contraction period. A similar result is reported by Granger (1996) using data from an international survey of forecasters. Loungani (2000) also reports that forecasters regularly miss recessions.

[1] The data set from which the time series were drawn is described in detail in Döpke and Fritsche (2004). We rely on annual data from 1963 to 2004. The growth forecast is the predicted growth rate of real GNP for the time span 1983 to 1989 and real GDP for all other years. The inflation forecast is the predicted change of the deflator of private consumption. As regards the actual outcome, we used the first published data. The numbers refer to West Germany up to 1992, and to the unified Germany from 1993 to present.

In Figure 1 we have plotted the forecast errors together with recession periods (shaded) in accordance with the business cycle and growth cycle concept of the Economic Cycle Research Institute (ECRI).[2] In almost all cases, large overestimations of growth coincide with recessions using either the traditional business cycle concept or the growth cycle concept. Underestimations of growth coincide with pre-recession periods where growth rates seem to be above average. With regard to inflation forecast errors, the relationship is unclear.

Figure 1: Forecast errors and recessions[3]



Regressions of the respective forecast errors on a constant and a dummy for the business cycle dates – for robustness check the dummies were defined according to mentioned different concepts – confirms the hypothesis of a relationship between forecast errors and recession periods.

[2] The business cycle concept refers to peaks and troughs in levels, whereas the growth cycle concept refers to peaks and troughs in the growth rate of GDP. Owing to the annual frequency used here, the dating can only be very imprecise. A year was counted as a recession year if more than 6 months were in the recession regime. For a more detailed description of the underlying business cycle dating procedure, refer to http://www.businesscycle.com.

[3] The business cycle dating stems from ECRI (www.businesscycle.com).

Table 1: Forecast errors and recessions: estimation results

Business cycle dummy defined according to …	**Estimations results (s.e. in parenthesis)**	
	Growth forecast errors	Inflation forecast errors
Business cycle concept	$e_t^g = \underset{(0.25)}{-0.20} + \underset{(0.60)^{***}}{1.79} D_t + \varepsilon_t$	$e_t^\pi = \underset{(0.26)}{0.06} - \underset{(0.43)}{0.51} D_t + \varepsilon_t$
Growth cycle concept	$e_t^g = \underset{(0.39)}{0.16} + \underset{(0.44)}{0.22} D_t + \varepsilon_t$	$e_t^\pi = \underset{(0.14)}{0.23} - \underset{(0.30)^{*}}{0.59} D_t + \varepsilon_t$

Note: The equations were estimated using Newey-West HAC standard errors & covariances (lag truncation = 3). *, **, *** denote significance at the 10, 5 and 1 per cent level respectively.

Table 1 shows that with regard to the growth forecasts and using the business cycle concept, the dummy is significant at 1 per cent and the coefficient has a positive value. According to this equation, the "six leading institutes" tended to overestimate GDP growth in recession periods (business cycle concept) by 1.6 percentage points on average. With regard to the inflation forecasts and using the growth cycle concept, the dummy is significant at 10 per cent and the coefficient has a negative value. According to this results, the "the six leading institutes" tended to underestimate inflation in growth cycle recession periods by –0.36 percentage points.

3 Testing for rationality

We start our empirical analysis with standard measures of forecasting accuracy and standard tests of forecast rationality. Table 2 presents some standard statistics.[4]

- The mean error (a positive (negative) value of the mean error corresponds to an under(over)estimation of the variable) is not significantly different from zero as can be seen from a regression on a constant.[5]
- The values of the mean absolute error as well as the root mean squared error show that forecast errors in general and over a longer time span are not small – a fact well known to professional observers but not necessarily in the public debate. The forecast error is about one third of the variance of the time series.

[4] Unless otherwise stated, our notation follows the textbook of Diebold (1998).

- We applied the normality test as described in Doornik and Hansen (1994). The hypothesis of normality is not supported for growth forecast errors. For inflation forecast errors, the normality hypothesis is supported.

Table 2: The track record of the forecasts

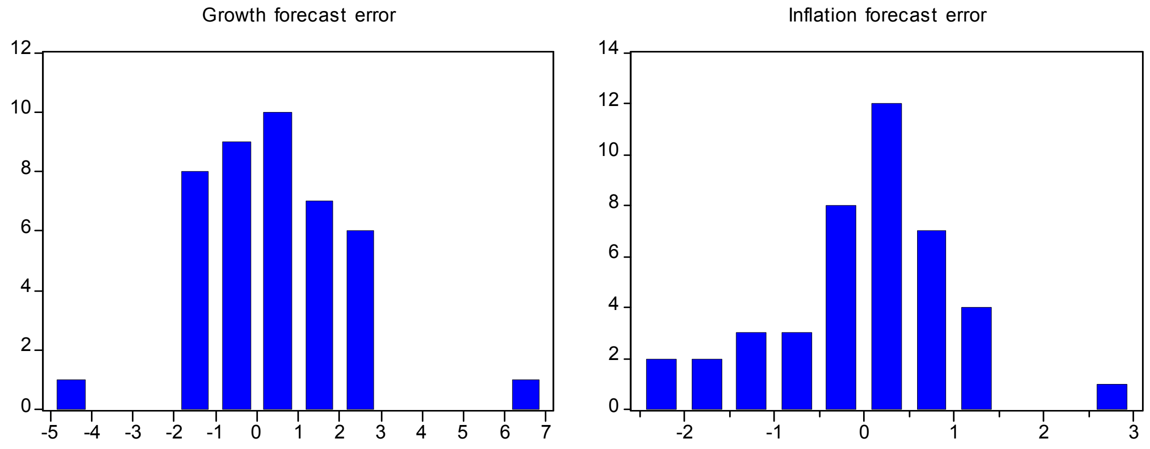
Statistic	Growth forecasts	Inflation forecasts
Mean error	0.27	-0.07
Mean absolute error	1.31	0.77
Root mean squared error	1.77	1.01
Regression of forecast error on a constant (p-value, Newey-West corrected)	0.33	0.78
RMSE as a percentage of variance of predicted series	0.35	0.29
Doornik-Hansen-Normality test (p-value)	0.004	0.36

Note: The first three rows are shown as percentage points

- Figure 2 shows the histograms for the growth and inflation forecast errors, respectively. The distribution of growth forecast errors is leptocurtic with a slightly longer right tail. Some "large" forecast errors dominate both outer tails. Inflation forecast errors are fairly normally distributed with a kurtosis slightly above 3 and a longer left tail which points to the tendency of forecasters to underestimate inflation.

[5] This refers to the overall sample.

Figure 2: Histograms



In the next step we turn to unbiasedness and efficiency/rationality tests.[6] Table 3 presents the respective results.

[6] In all cases, we checked for absence of autocorrelation and for homoscedasticity since the respective test statistics are only valid under these assumptions. The condition is fulfilled for all growth specifications but not for inflation specifications. We therefore used a Newey-West correction. The detailed results of Breusch-Godfrey LM tests and White heteroscedasticity tests are not reported here but are available from the authors on request.

Table 3: Results of rationality tests

Statistic	Growth forecasts		Inflation forecasts	
	OLS	LAD	OLS	LAD
Test for unbiasedness (p-value)	0.33	1.00	0.69	0.22
Test for weak rationality (p-value)	0.75	0.54	0.001	0.001
i) Test for strong rationality based on short-term interest rates (Wald test, p-value)	0.10	0.06	0.33	0.50
ii) Test for strong rationality based on short-term real interest rates (Wald test, p-value)	0.61	0.32	0.32	0.58
iii) Test for strong rationality based on interest rate spread (Wald test, p-value)	0.15	0.04	0.31	0.41
iv) Test for strong rationality based on the change in real external value of the currency (Wald test, p-value)	0.45	0.89	0.24	0.44
v) Test for strong rationality based on the change in the oil price (Wald test, p-value)	0.22	0.67	0.33	0.57
vi) Test for strong rationality based on the change in OECD industrial production (Wald test, p-value)	0.57	0.55	0.05	0.04
vii) Test for strong rationality based on the monetary stance indicator (Wald test, p-value)	0.01	0.02	0.23	0.28

- First, we test for the unbiasedness of the forecasts. To this end, it is possible to make use of the Mincer-Zarnowitz regression, i.e. running the following regression: $y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1,t} + u_t$ (where y_{t+1} is the "actual" in period t+1 and $\hat{y}_{t+1|t}$ is the forecast made in period t for period t+1) and testing the hypothesis $H_0 : \beta_0 = 0 \wedge \beta_1 = 1$. The relevant column in the table reports the p-value of this test under the header "OLS". Both forecasts are classified as unbiased.
- As a test for the forecasts' rationality we extend the Mincer-Zarnowitz regression by exogenous variables, i.e. estimating the equation $y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1,t} + \beta_2 X_{t-1} + u_t$ by OLS where X_{t-1} represents a respective exogenous variable. We use a standard Wald test to check the hypothesis $H_0 : \beta_0 = 0 \wedge \beta_1 = 1 \wedge \beta_2 = 0$ (Holden and Peel, 1990). If this hypothesis cannot be rejected, the forecast has to be considered as rational. We test a weak and a strong version of this test.

- The "weak" version refers to a test for autocorrelation of the forecast errors. For an optimal forecast, one should be unable to find any variable which helps to forecast the errors. As a consequence, the lagged forecast errors should also be non-informative for the ex ante error. In the literature, this is sometimes referred to as a test for "weak efficiency" or "weak rationality" (Kirchgässner, 1984). We define $y_{t+1} - \hat{y}_{t+1,t} = e_{t+1}$ and estimate $y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1,t} + \beta_2 e_t + u_t$. We report the p-value of the Wald test that $H_0 : \beta_0 = 0 \wedge \beta_1 = 1 \wedge \beta_2 = 0$ under the header "OLS". The test is unable to reject the null of an efficient use of the available information for growth forecasts but not for inflation forecasts. However, one should keep in mind that the information set represented by the lagged forecast error is rather limited.

- A stronger test version of the hypothesis of efficient information processing – sometimes referred as "rational expectation formation" or "strong rationality" – stipulates that the forecast errors are uncorrelated to *any* variables known to the forecasters at the time of the forecast. Using these kind of arguments Hagen and Kirchgässner (2001), Kirchgässner and Savioz (2001) as well as Harvey (1991) find evidence that forecasting equations based e.g. on monetary variables outperform the institutes' forecasts. This implies that the latter could not be efficient. However, as it is emphasised by Tichy (1994), tests which rely on exogenous variables are problematic since it is unclear what the forecasters *exactly* know about the future stance of the cycle. Therefore we selected a number of macroeconomic variables from the second quarter of the year in which the forecast is made and repeat the test with exogenous variables instead of the lagged forecast errors as explanatory variables. If the hypothesis of the above-mentioned Wald test cannot be rejected, the forecast has to be considered as rational in a strong sense. The results (estimated by OLS) are reported under the header "OLS".In particular, we use the following exogenous variables:
  i) Short-term interest rates as of the second quarter of the forecasting year as a proxy for the stance of monetary policy. [7]
  ii) Real short-term interest rates as of the second quarter of the forecasting year as a further proxy for the stance of monetary policy.[8]

[7] We use 3-month FIBOR/EURIBOR as a measure of short-term interest rates.

iii) Interest rate spread as of the second quarter of the forecasting year as a potentially good leading indicator (Estrella, Rodrigues and Schich 2003).[9]

iv) The change over previous year as of the second quarter of the forecasting year of the real external value of the domestic currency to capture possible exchange rate shocks.

v) The change over previous year as of the second quarter of the forecasting year of the oil price as a proxy for supply side shocks.

vi) The change over previous year as of the second quarter of the forecasting year of industrial production in all OECD countries to take into account demand fluctuations outside Germany.

vii) A monetary policy "stance" indicator constructed along the lines proposed by St-Amant (1996) and Gottschalk (2001).[10] The indicator measures the deviation of the real short-term interest rate from its model-based equilibrium level.

The results of the strong rationality tests indicate that rationality is sometimes rejected – more often for growth than for inflation forecasts. It is particularly puzzling that past short-term interest rates, the interest rate spread or the monetary policy indicator of the respective second quarter of the forecasting year – information which is definitely known to all the involved forecasters – help to improve the fit of the equation and, thus, violate the rationality of the forecasts (see Kirchgässner and Savioz, 2001): Even a brief examination of the joint forecast publication will make apparent that virtually all forecasters agree on the importance of monetary policy for the business cycle. As a consequence, the forecasters closely monitor short-term interest rates. Since underestimations of growth are strongly related to recession periods, this points to the possibility that forecasters tend to underestimate the dampening effects of some macroeconomic variables – notably changes in the stance of monetary policy or external demand.

---

[8] The real interest rate is the short-term interest rate deflated by CPI inflation.

[9] The term spread is defined as "Umlaufsrendite" minus 3-month FIBOR/EURIBOR.

[10] We use a structural VAR approach to disentangle expected real interest rates into that part driven by monetary policy and that part reflecting other sources. The identification scheme relies on the applicability of the Fisher equation and uses long-run restrictions as in Blanchard and Quah (1989).

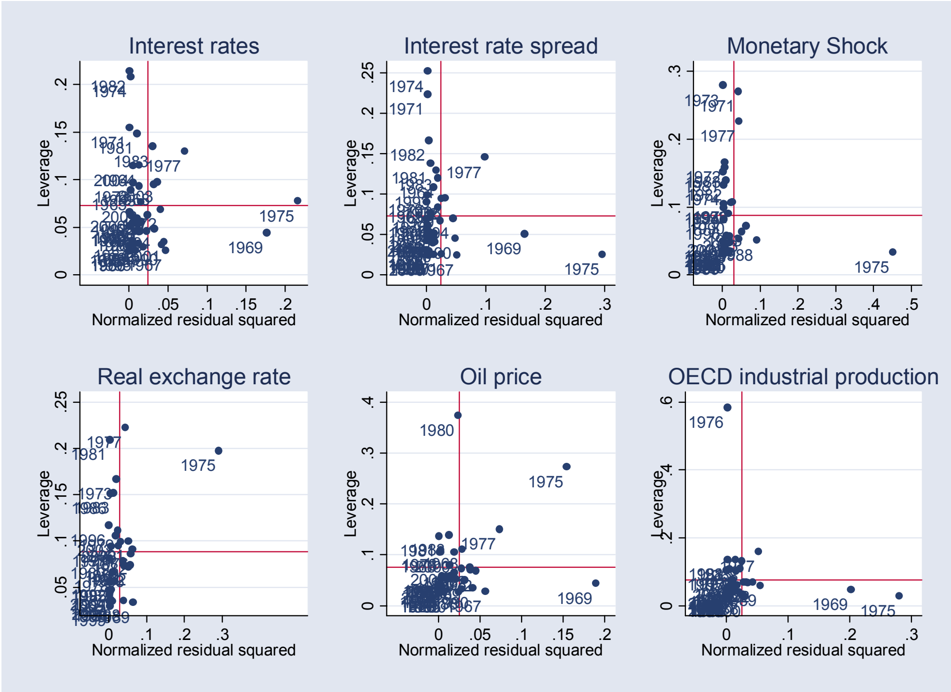
The often-stated hypothesis that the alleged poor quality of Germany's business cycle forecasts is due to external shocks is, on the other hand, not supported – at least not by this test.

Turning to inflation forecasts, the OECD industrial production has some explanatory power for the inflation forecast error. External shocks seem to matter for inflation forecast errors rather than growth forecast errors. All in all, our results give support to doubts concerning the rationality of forecasts.

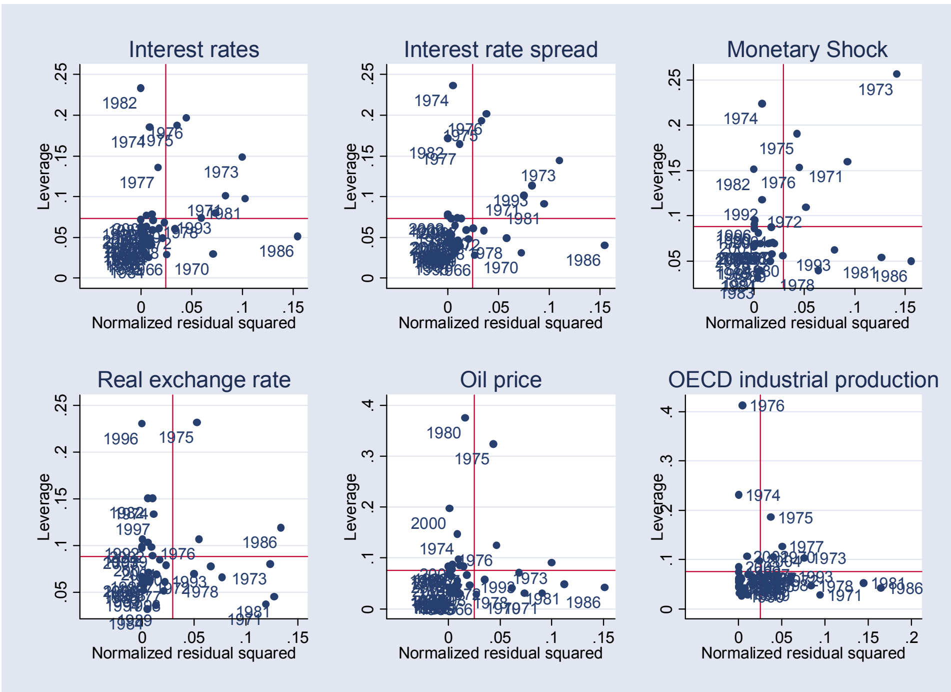
To gain further evidence on potential reasons for the rejection of the hypothesis of strong rational expectations, it might be useful to consider whether single influential outliers are responsible for this result. Thus, we employ techniques of outlier detection and robust estimation techniques. In particular, we first take a look at the leverage-versus-squared-residuals plot. A point above the horizontal line shows years with a high leverage, ie years with a particular influence on the estimated regression. Years with point far to the right represent years with a high residual. It is possible to drop the outliers from the data. If these observations can reasonably be assumed to be valid within the respective model under consideration, it makes more sense to re-run the regressions testing for strong rationality using a robust estimator (Greene, 2003 ch. 16). A possible estimator is given by the least absolute deviations estimator (LAD).[11] We suspect that the rejection of the hypothesis of strong rationality depends of a limited number of influential observations. If this is true, the robust estimator should not reject the hypothesis of rationality in those cases where the OLS based results lead to a rejection of the hypothesis. The respective robust results are reported under the header "LAD" in table 3.

[11] LAD is consistent and asymptotically normal under quite broad conditions.

Figure 3: Outlier detection in rationality tests



Panel (a): GDP forecasts



Panel (b) Inflation forecasts

To begin with, figure 3, panels(a) and (b) depict the leverage-versus-squared-residuals plots for the Holden/Peel-type test for forecast rationality. Visual inspection of the exhibits makes clear that, depending on the exogenous variable included in the test, some data points have a relatively high leverage, i.e. influence the regression more than others. The figures reveal that regarding the growth forecasts some data points have a particularly large influence on the estimated relation. In case of all monetary variables, years frequently seen as periods of marked monetary tightening (1971, 1974 and 1982) show a large leverage. As regards the real exchange rate, the value for the year 1977 appears to be important for the estimated coefficients. While one would have expected the observation of the year 1980 in case of the oil price, since this years marks the beginning of the second oil price crisis, the fact that the data points for 1973 or 1974 show not similar influence comes as a surprise. Looking at the industrial production abroad, 1976 is the observation with the strongest influence on the coefficients and, thus, on the test result for rationality. In most cases the observation in 1975 is an outlier, that is, the year represents a huge miss-prediction, but with no strong influence on the estimation results.

Regarding the inflation forecasts, the picture is less systematic. The year 1974 appears to be influential in all equation with monetary variables as regressors, but several other years also appear to be important with no obvious systematic pattern. For all equations, 1996 appears to be a large, though not influential outlier. In this year, the raw material prices witnessed an marked drop.

The results for the rationality tests based on robust estimation methods show no remarkable difference to the OLS results. From that perspective, influential outliers do not seem to be the prime suspect as a reason for the rejection of rationality in several cases.

4 Testing for non-linearities

The results of the previous parts of this paper suggest that the investigated macroeconomic forecasts for Germany did not pass tests on strong rationality, namely the test of whether exogenous variables could improve the forecasts. In this section, we aim to gain further insights into the nature of the relationship between forecast errors and the state of the economy. In particular, we test for a possible non-linear relationship between lagged macroeconomic variables and forecast errors.

Consider, for example, the test for rationality by running regressions of the form $y_{t+1} = \beta_0 + \beta_1 \hat{y}_{t+1|t} + \beta_2 X_{t-1} + u_t$. To simplify the analysis, we restrict $\beta_1 = 1$ - which is satisfied because the tests in section 2 revealed, that the unbiasedness condition is fulfilled - and test the equation $y_{t+1} - \hat{y}_{t+1|t} = e_{t+1} = \beta_0 + \beta_2 X_{t-1} + u_t$ where X_{t-1} stands for all the above-mentioned variables.

With this equation at hand, we tested for non-linearity by applying a Ramsey (1969) RESET test with power 2 and 3 and a Tsay (1989) arranged regression test. Furthermore, we used a threshold estimation as described in Hansen (1997) with bootstrapped critical values to be aware of small sample problems.[12]

The results given in Table 4 indicate only one (and quite weak) non-linear relationship between growth forecast errors and a macroeconomic variable, namely the interest rate spread. The significance level of the Tsay (1989) test delivers a p-value of 0.11 and the Hansen (1997) test gives a p-value of 0.14. There are no signs of non-linearity between growth forecast errors and any other indicator.[13] There are however signs of a non-linear relationship between inflation forecast errors and several macroeconomic indicators. As Table 5 indicates, there are signs of a non-linear relationship between inflation forecast errors on the one hand and short-term interest rates, interest rate spread (weak evidence, p-value for Tsay (1989) test: 0.11, p-value for Hansen (1997) test: 0.14) and the monetary policy stance indicator on the other hand.

[12] The calculations were made using RATS 5.04 and the procedures RESET.SRC, TSAYTEST.SRC, THRESHOLD.SRC available at http://www.estima.com. The critical values for the Hansen (1997) test were obtained by using the bootstrapping procedure embedded in THRESHOLD.SRC using 5,000 repetitions.

[13] Note, however, that the threshold procedure trims the range of the ordered time series by 15 per cent from each side. Especially for the oil price, there is some evidence that the relationship between oil price changes and GDP growth is highly non-linear and dominated by very large spikes in the oil price. Owing to the limitation of data in our case, we were unable to use more sophisticated methods. See Hamilton (2003).

Table 4: Non-linearity tests: growth forecast errors

	RESET test (Ramsey, 1969)				Tsay (1989) arranged regression test		Hansen (1997) test for a break in a linear regression based upon a threshold variable		
	H_0: No departure from linearity in the conditional mean				H_0: No threshold against SETAR model		H_0: No threshold against alternative of threshold under maintained assumption of homoskedastic errors		
	F-test stat. / p-value				F-test stat. / p-value		F-test stat. / p-value *		threshold
	power = 2		power = 3						
Short-term interest rate	0.61	0.44	0.31	0.74	0.43	0.65	3.11	0.41	3.61
Short-term real interest rate	0.19	0.66	0.23	0.79	0.47	0.63	0.92	0.97	1.01
Interest rate spread	0.09	0.77	0.06	0.94	0.11	0.89	4.68	0.14	-1.20
Change in oil price	1.38	0.24	1.70	0.19	0.60	0.55	1.65	0.80	-0.33
Monetary policy indicator (SVAR)	0.17	0.67	0.54	0.58	0.83	0.44	1.35	0.92	-0.67
Change in real effective exchange rate	1.98	0.17	1.07	0.35	1.25	0.30	2.30	0.69	0.034
Change in OECD industrial production	0.06	0.80	1.13	0.33	0.27	0.76	2.69	0.51	-0.014

Note: * The p-value was calculated using 5,000 bootstrap replications.

Table 5: Non-linearity tests: inflation forecast errors

	RESET test (Ramsey, 1969)				Tsay (1989) arranged autoregression test		Hansen (1997) test for a break in a linear regression based upon a threshold variable		
	H0: No departure from linearity in the conditional mean				H0: No threshold against SETAR model		H0: No threshold against alternative of threshold under maintained assumption of homoskedastic errors		
	F-test stat. | p-value				F-test stat. | p-value		F-test stat. | p-value *		threshold
	power = 2		power = 3						
Short-term interest rate	0.75	0.39	0.37	0.69	0.69	0.51	6.40	0.04	4.11
Short-term real interest rate	1.01	0.32	1.30	0.29	0.96	0.39	2.54	0.55	-1.03
Interest rate spread	2.08	0.15	1.23	0.230	2.03	0.14	5.06	0.11	0.26
Monetary policy indicator (SVAR)	16.33	0.0003	8.35	0.001	8.72	0.001	10.88	0.004	-0.26
Change in oil price	0.08	0.77	0.90	0.41	0.49	0.61	1.44	0.86	-0.14
Change in real effective exchange rate	0.07	0.79	0.42	0.65	0.64	0.53	2.31	0.69	-0.045
Change in OECD industrial production	0.24	0.62	0.53	0.59	0.06	0.95	1.61	0.83	0.0007

Note: * The p-value was calculated using 5,000 bootstrap replications.

For those indicators, where non-linearity is indicated by the threshold test (including the interest rate spread, where the test results were not very decisive), we estimated a two-regime threshold model of the following form:[14]

$$e_t = I_t[\alpha_0 + \alpha_1 X_{t-1}] + (1 - I_t)[\beta_0 + \beta_1 X_{t-1}] + \varepsilon_t$$
$$I_t = \begin{cases} 1 \text{ if } X_{t-1} \geq \tau \\ 0 \text{ if } X_{t-1} < \tau \end{cases}$$

The results are given in table 6 and a graphical analysis is given in figures 4 to 6.

Table 6: Threshold models: estimation results

Variable	**Estimations results (s.e. in parenthesis)**	**Estimated Threshold (τ)**
	Growth forecast error	
Model I: Interest rate spread	$e_t^g = I_t\left[+\underset{(0.36)}{0.45} - \underset{(0.32)}{0.51} \cdot X_t\right] + (1 - I_t)\left[-\underset{(2.11)}{3.17} - \underset{(0.69)^{*}}{1.36} \cdot X_t\right] + \varepsilon_t$	-1.20
	Inflation forecast error	
Model II: Interest rate spread	$e_t^\pi = I_t\left[-\underset{(0.41)}{0.47} + \underset{(0.31)}{0.21} \cdot X_t\right] + (1 - I_t)\left[+\underset{(0.27)^{**}}{0.67} + \underset{(0.14)^{***}}{0.41} \cdot X_t\right] + \varepsilon_t$	0.26
Model III: Monetary policy indicator (SVAR)	$e_t^\pi = I_t\left[+\underset{(0.26)}{0.36} - \underset{(0.23)^{**}}{0.57} \cdot X_t\right] + (1 - I_t)\left[+\underset{(0.42)^{***}}{1.48} + \underset{(0.31)^{***}}{1.15} \cdot X_t\right] + \varepsilon_t$	-0.26

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

To begin with, for both growth and inflation forecast errors, we found reasonable interpretions for the model with the interest rate spread. As the estimation outcome for model I shows, there is a regime $(1 - I_t)$ connected with a negative interest rate spread (below –1.20) which coincides with relatively high negative growth forecast errors and another regime where the errors are minor – see figure 4 for a graphical representation as well. The result is consistent with the literature on the leading indicator properties of the interest rate spread with regard to recessions (Estrella, Rodrigues and Schich 2003), which is probably not taken into account in a proper way when making the forecast.

[14] We used the program for grid search as described in Enders (2003).

Figure 4: Regimes in Threshold Model I: Growth forecast error and interest rate spread

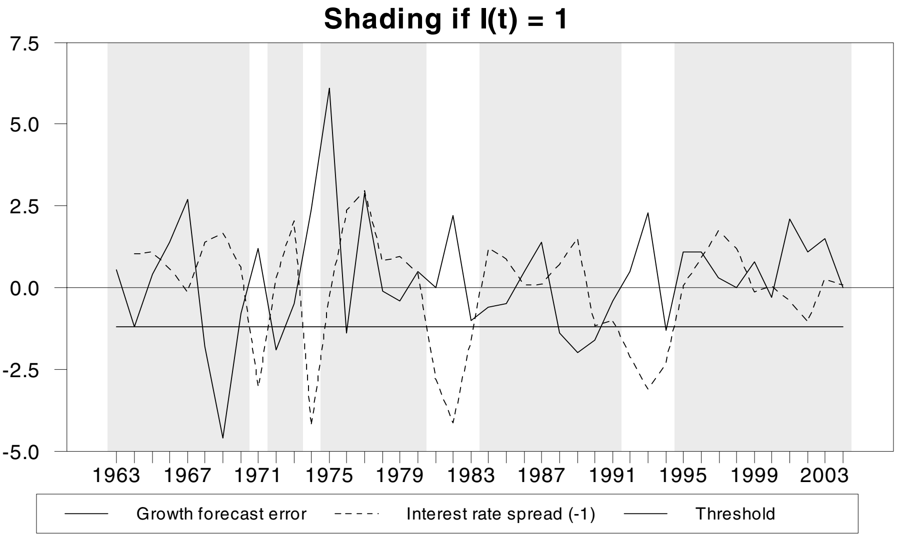


Figure 5: Regimes in Threshold Model II: Inflation forecast error and interest rate spread

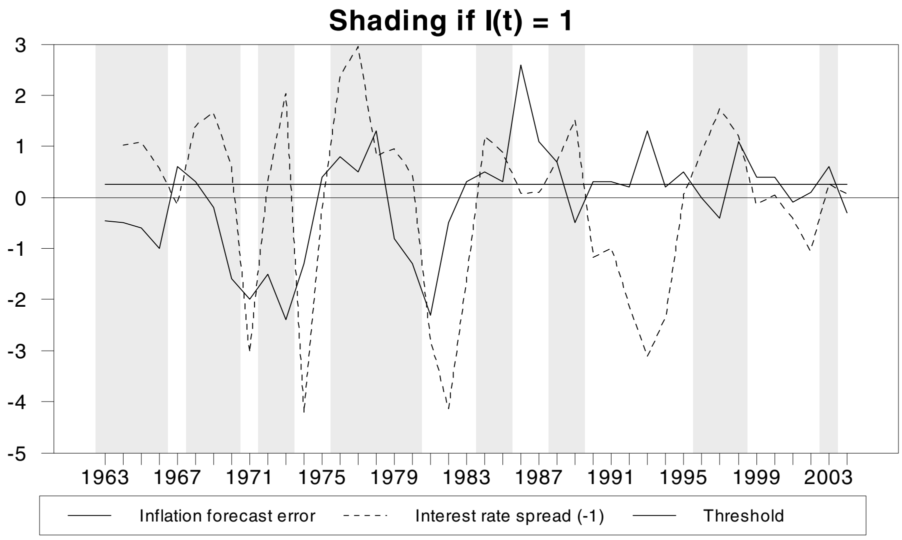
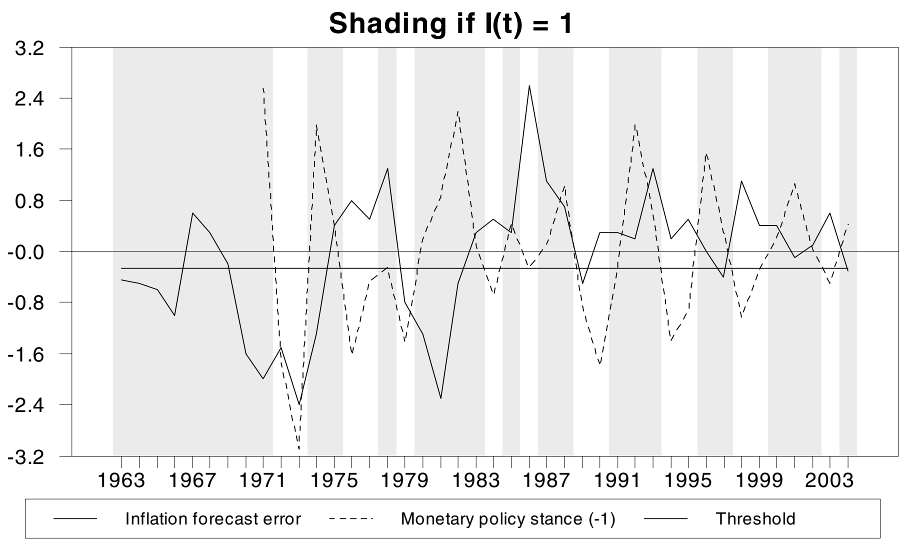


Figure 6: Regimes in Threshold Model III: Inflation forecast error and monetary policy indicator



The results for model II shows that the interest rate spread might also be useful for explaining inflation forecast errors when the interest rate spread falls below 0.26. The coefficients for regime where I_t is equal to 0 turn out to be significant, whereas the coefficients of the regime indicated by $I_t = 1$ are statistically not significant. Figure 5 shows the graphical representation. The third estimated model for the relation between inflation forecast errors and the monetary policy stance indicator (model III) also indicates a statistically significant non-linear relationship between these variables. It relates mainly positive inflation forecast errors (inflation underestimations) to periods of stimulating monetary policy (regime $I_t = 0$ which indicates periods of a monetary policy stance below a slightly negative threshold of –0.26). This result can also be seen when looking at Figure 6.[15]

5 What explains "large" forecast errors?

In the next step, we would like to differentiate between "large" and "small" overestimations and underestimations of growth and inflation. Taking a closer look at "large" forecast errors

[15] From the graphical representation it becomes obvious that the errors became smaller over time. The results might therefore be subject to structural breaks. We tested for structural breaks within the threshold models using the Hansen (1991) procedure available in RATS (STABTEST.SRC at www.estima.com). No structural break for the overall equation was detected in any case.

might be very helpful in further understanding the reasons for errors. This makes it necessary to define the notion "large forecast errors" more precisely. We calculated the quartiles of the distribution of the respective forecast errors and count as "large" forecast errors the data points which are in the outer quartiles of the distribution on both tails. Errors in the first quartile were counted as "large" underestimations, errors in the fourth quartile as "large" overestimations.

The probit approach as used by Estrella and Mishkin (1998) is then used to determine whether a variable helps to explain "large" forecast errors. The dependent variable utilised in this analysis is a dummy that takes the value 1 for a large error (outside interquartile range) and 0 otherwise. Thus, the following equation is estimated:

$$\Pr\left(E_t = 1 \middle| X_{t-1}, \beta\right) = 1 - \Phi(-(X_{t-1})'\beta)$$

where E_t is a dummy for the "large" forecast error, Φ is the cumulative distribution function of the standard normal distribution, and X_{t-1} is the indicator to be considered. E_t takes the form

$$E_t = \begin{cases} 1 & \text{if a large forecast error occurs} \\ 0 & \text{else} \end{cases}$$

As indicator variables for X_{t-1} which may help to explain forecast errors, we considered all the macroeconomic variables as described above.

Table 7: Probit models: growth forecast errors

Variable	Constant	p-value constant	β_1	p-value β_1	Pseudo R^2
Short term interest rate	-1.61	0.00	0.15	0.06	0.08
Real interest rate	-1.05	0.01	0.12	0.30	0.02
Interest rate spread	-0.46	0.08	-0.20	0.10	0.06
Oil price changes	-0.76	0.00	0.00	0.48	0.01
Monetary policy indicator	-0.80	0.00	0.40	0.07	0.10
Real exchange rate changes	-0.55	0.89	0.00	0.96	0.00
Changes in OECD industrial production	-0.88	0.00	0.05	0.41	0.02

Panel (a): Large overestimations

Variable	Constant	p-value constant	β_1	p-value β_1	Pseudo R^2
Short term interest rate	-0.49	0.35	-0.04	0.64	0.00
Real interest rate	-0.48	0.20	-0.09	0.47	0.01
Interest rate spread	-0.90	0.00	0.12	0.37	0.02
Monetary policy indicator	-0.96	0.00	-0.47	0.04	0.14
Oil price changes	-0.66	0.00	-0.01	0.38	0.02
Real exchange rate changes	-0.80	0.00	0.00	0.99	0.01
Changes in OECD industrial production	-0.46	0.09	-0.09	0.13	0.06

Panel (b): Large underestimations

The interpretation of the results is complicated slightly by the fact that the coefficients cannot directly be interpreted as the marginal effect on the dependent variable. The effect of a small change of X_{t-1} on the probability of making a "large" error , however, can be calculated from the estimated coefficients and a table for a cumulative standard normal distribution.

We report the estimation results for overestimations and underestimations of growth and inflation forecasts respectively. For those indicators where we found significant effects, the effects of a variation of X_{t-1} on $\Pr(E_t)$ were plotted using probability response curves.

Table 8: Probit models: inflation forecast errors

Variable	Constant	p-value constant	β_1	p-value β_1	Pseudo R^2
Short term interest rate	-0.42	0.44	-0.07	0.46	0.01
Real interest rate	-0.62	0.11	-0.07	0.58	0.01
Interest rate spread	-0.91	0.00	0.08	0.56	0.01
Monetary policy indicator	-0.73	0.00	-0.12	0.54	0.01
Oil price changes	-0.74	0.00	-0.01	0.33	0.03
Real exchange rate changes	-0.70	0.00	0.01	0.87	0.02
Changes in OECD industrial production	-0.56	0.04	-0.08	0.16	0.05

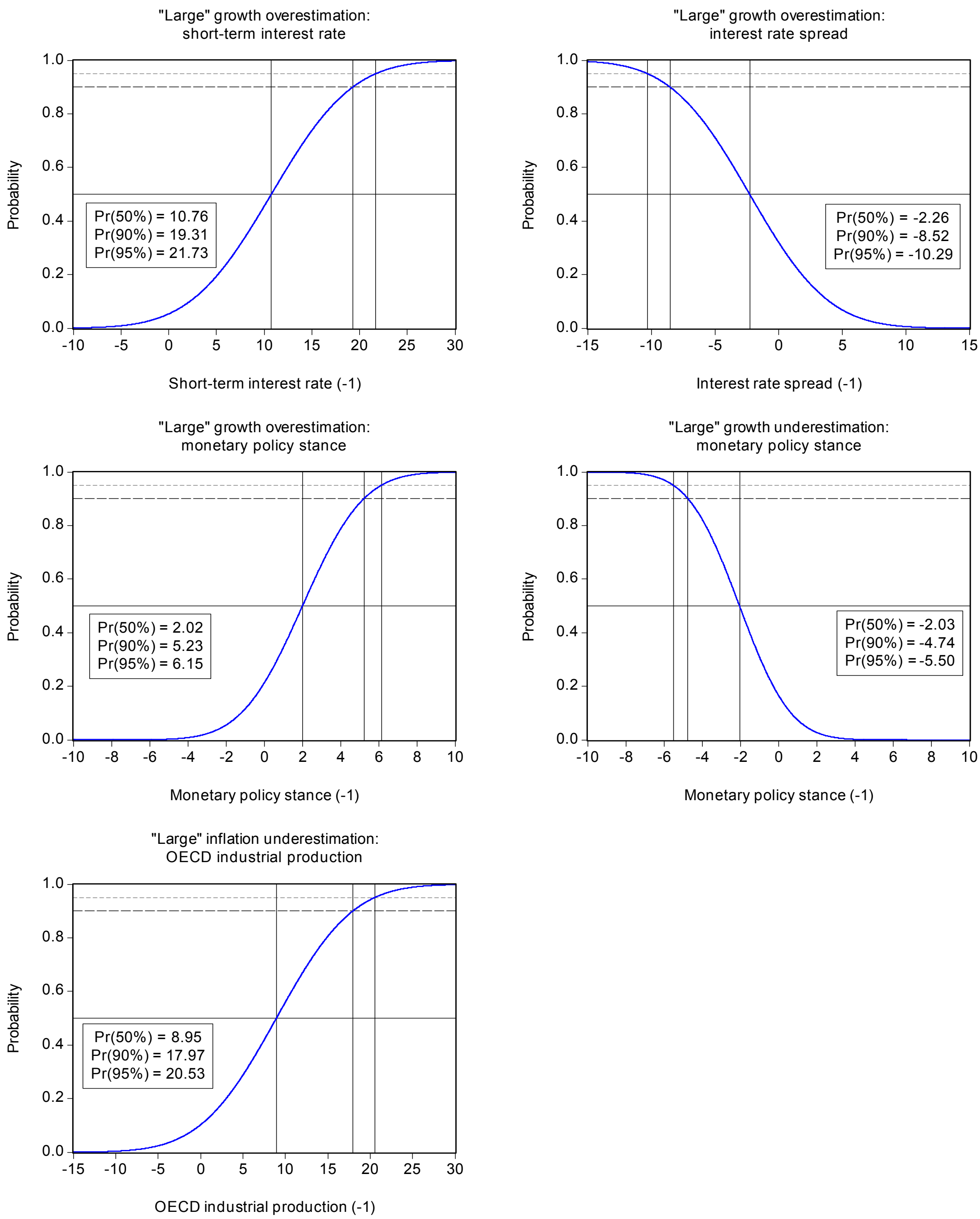
Panel (a): Large overestimations

Variable	Constant	p-value constant	β_1	p-value β_1	Pseudo R^2
Short term interest rate	-1.26	0.02	0.09	0.24	0.03
Real interest rate	-0.69	0.06	-0.01	0.95	0.00
Interest rate spread	-0.58	0.03	-0.10	0.43	0.01
Monetary policy indicator	-0.82	0.00	-0.03	0.86	0.00
Oil price changes	-0.80	0.00	0.01	0.25	0.03
Real exchange rate changes	-0.94	0.00	0.06	0.40	0.12
Changes in OECD industrial production	-1.27	0.00	0.14	0.05	0.10

Panel (b): Large underestimations

We find that the short-term interest rate as well as the interest rate spread and the monetary policy indicator are related to large growth overestimations. This is in line with the literature on monetary policy effects, which states that periods of high interest rates seem to be connected with unexpected recessions (see Ball and Croushore, 2003; Romer and Romer, 1989, 2003) and therefore – as described in section II - "large" forecast errors. The monetary policy indicator is also related to "large" growth underestimations which indicates that monetary policy might be effective in both directions. Furthermore, we find that changes in the OECD industrial production are somehow related to "large" inflation errors – especially underestimations.

Figure 7: Probit models: probability response curves



Having the probability responses at hand, it can be analysed at which value of X_{t-1} the probability of making a "large" error reaches several values. The figures show e.g. that a monetary policy stance indicator (the deviation of the short-term real interest rate from its model-based equilibrium) above 2.02 would indicate a more than 50 % probability of a "large" growth overestimation, a value above 5.23 would indicate a more than 90 % chance of making a "large" growth overestimation and a value above 6.15 would indicate a more than 95 % chance of making a "large" growth overestimation. The respective threshold values (50 %, 90 %, 95 %) for all other variables are also reported.

6 Conclusion

The paper analyses the reasons for departures from strong rationality of German business cycle forecasts based on annual observations form 1963 to 2004. To this end, we rely on forecasts from the joint forecast of the so-called "six leading" forecasting institutions. Both growth and inflation forecasts are investigated.

We document a relationship between forecast errors and recessions. We also find that macroeconomic variables known at the period of the forecast are informative with regard to the forecast errors. This in turn leads to the rejection of rationality in the strong sense in some cases. We suspect that the relation between forecast errors and macroeconomic fundamentals is non-linear. Using several tests and models, some evidence for such a non-linearity is found. Based on threshold models, we show how the interest rate spread and monetary policy stance changes are related to growth and inflation forecast errors in a non-linear way. To get deeper insights we distinguish between over-and underestimations and introduce the concept of "large" errors. Evidence from probit models further supports the notion that some macroeconomic fundamentals correlate to "large" overestimations and underestimations. Monetary disturbances seem to be important for "large" growth over- and underestimations.

All in all, we conclude that some macroeconomic variables – especially monetary factors – which are known at the date of the forecast are related to forecast errors in a non-linear way. It raises doubts that professional forecasters are fully rational in that they always take into account the "appropriate" model of the economy. Maybe they do not even know the stance of the economy or maybe the model the forcasters have in mind is "wrong". Monetary policy might then have real effects just because it is unexpected by professional forecasters and people rely on professional forecasters' results.

7 References

Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute (various issues). Zur Lage der Weltwirtschaft und der deutschen Wirtschaft, Berlin.

Ash, J.C.K., *Smyth*, D.J. and *Heravi*, S.M. (1998). Are OECD Forecasts Rational and Useful?: A Directional Analysis, International Journal of Forecasting, 14, 381-391.

Ball, L. and D.D. *Croushore* (2003). Expectations and the Effect of Monetary Policy. Journal of Money, Credit and Banking 35: 473-484.

Blanchard, O. and *D. Quah* (1989). The dynamic effects of aggregate demand and supply side disturbances. American Economic Review 79: 655-673.

Carroll, C.D. (2003). Macroeconomic expectations of households and professional forecasters. Quarterly Journal of Economics 118: 269-298.

Diebold, F.X. and *J.A. Lopez* (1996). Forecast Evaluation and Combination. In: Maddala, G.S. und C.R. Rao (ed.): Handbook of Statistics Vol. 14: 241-268.

Doornik, J.A. and *H. Hansen* (1994). A Practical Test for Univariate and Multivariate Normality. Discussion Paper, Nuffield College.

Döpke, J. and *U. Fritsch*e (2004). Growth and inflation forecasts in Germany: an assessment of accuracy and dispersion. DIW Discussion Paper No. 399. Berlin.

Döpke, J. and *Langfeldt, E.* (1995). Zur Qualität von Konjunkturprognosen für Westdeutschland 1976-1994. Kiel Discussion Papers No. 274. Kiel Institute of World Economics, Kiel.

Elliot, G, I. Komunjer and *A. Timmermann* (2004). Biases in Macroeconomic Forecasts: Irrationality or Asymmetric Loss?, mimeo, http://www.hss.caltech.edu/~komunjer/respaper.htm

Enders, W. (2003). RATS Programming Manual. http://www.estima.com/enders/index.shtml

Estrella, A., A.P. Rodrigues and *S. Schich* (2003). How Stable is the Predictive Power of the Yield Curve? Evidence from Germany and the United States. Review of Economics and Statistics 85: 629-644.

Fintzen, D. and *Stekler, H.O.* (1999). Why Did Forecasters Fail to Predict the 1990 Recession?, International Journal of Forecasting, 15, 309-323.

Fildes, R. and *Stekler, H.O*. (2002). The state of macroeconomic forecasting. Journal of macroeconomics 24: 435-468.

Gottschalk, J. (2001). Measuring expected inflation and the ex-ante real interest rate in the euro area using structural vector autoregressions. Kiel Working Paper No. 1021. http://opus.zbw-kiel.de/volltexte/2003/71/

Granger, C.W.J. (1996). Can we Improve the Perceived Quality of Economic Forecasts? Journal of Applied Econometrics, 11, 455-473.

Green, W.H. (2003). Econometric Analysis.New Jersey: Prentice Hall.

Hagen, H.M. and *Kirchgässner, G.* (1997). Interest Rate Based Forecasts vs. Professional Forecast for German Economic Growth: A Comparison. Discussion Paper No. 9714. University of St. Gallen.

Hamilton, J. D. (2003). What Is an Oil Shock?, Journal of Econometrics, April 2003, vol. 113, pp. 363-398.

Hansen, B. (1991). Parameter Instability in Linear Models, Journal of Policy Modeling, 14 (4), 1991, pp. 517-533.

Hansen, B. (1997). Inference in TAR models. Studies in non-linear Dynamics and Econometrics 2, Issue 1, Article 1.

Harvey, C.R. (1991). Interest Rate Based Forecasts of German Economic Growth, Weltwirtschaftliches Archiv, 127, 701-718.

Heilemann, U. (1998). Paradigm lost - Zu den Projektionen des Jahreswirtschaftsberichts der Bundesregierung. In: U. Heilemann, D. Kath, and N. Kloten (Eds.), Entgrenzung als Erkenntnis- und Gestaltungsaufgabe. Festschrift für Reimut Jochimsen zum 65. Geburtstag. Berlin: Duncker and Humblot, pp. 79-100.

Helmstädter, E. (1989). Zur Autokorrelation von Konjunkturprognosen, ifo-Studien, 35, 158-164.

Hinze, J. (1996). Konjunkturprognosen unter veränderten Rahmenbedingungen: Auswirkungen der deutschen und europäischen Einigung, Hamburger Jahrbuch für Wirtschafts- und Gesellschaftspolitik, 41, 59-84.

Holden, K. and *Peel, D.A.* (1990). On Testing for Unbiasedness and Efficiency of Forecasts, The Manchester School, 18, 120-127.

Kirchgässner, G. (1991). Sind die Prognosefehler der Konjunkturprognosen autokorreliert?, ifo-Studien, 37, 151-158.

Kirchgässner, G. and *M. Savioz* (2001). Monetary Policy and Forecasts for Real GDP growth: an Empirical Investigation for the Federal Republic of Germany. Germany Economic Review 2: 339-365.

Laster, D., P. Bennet and *I.S. Geoum* (1999). Rational Bias in Macroeconomic Forecasts. Quarterly Journal of Economics 114: 293-318.

Loungani, P. (2000). How accurate are private sector forecasts?: Cross-country evidence from consensus forecasts of output growth. International journal of forecasting. Vol. 17 (2001), No. 3, pp. 419-432.

Nierhaus, W. (1998). Ifo-Konjunkturprognosen und Wirklichkeit - eine Ergebnisanalyse für die Bundesrepublik im Zeitraum 1991–1997, ifo-Wirtschafts-Konjunktur, 50, 1-11.

Mankiw, G. Reis and *J. Wolfers* (2003). Disagreement about inflation expectations. NBER macroeconomics annual 18: 209-248.

Ramsey, J. B. (1969). Tests for Specification Errors in Classical Linear Least Squares Regression Analysis, Journal of the Royal Statistical Society, Series B, 31, 350–371.

Romer, Ch. D and *Romer, D. H.* (1989). Does monetary policy matter?: a new test in the spirit of Friedman and Schwartz. NBER macroeconomics annual. 121-170.

Romer, Ch. D. and *Romer, D. H.* (2003). A new measure of monetary shocks: derivation and implications. Cambridge, Mass. (= NBER Working Paper. 9866).

St-Amant, P. (1996). Decomposing U.S. Nominal interest rates into expected inflation and ex ante real interest rates using structural VAR methodology. Bank of Canada Working Paper 96-2.

Stäglin, R. (1998). Beiträge der Wirtschaftsforschungsinstitute zur kurzfristigen Wirtschaftsbeobachtung. Allgemeines Statistisches Archiv 82: 66-80.

Tsay, R.S. (1989).Testing and Modeling Threshold Autoregressive Processes, Journal of the American Statistical Asccosiation 84: 231-240.

Tsay, R.S. (2002) Nonlinear Models and Forecasting, in: M.P.Clements and D.F. Hendry (ed.), A Companion to Economic Forecasting, Oxford: 453-484

Stäglin, R. (1998). Beiträge der Wirtschaftsforschungsinstitute zur kurzfristigen Wirtschaftsbeobachtung. Allgemeines Statistisches Archiv 82: 66-80.

van Suntum, U. (1991). Konjunkturprognosen: Besser als ihr Ruf. Wirtschaftsdienst, 365-368.