Monetary Policy and Forecasts for Real GDP Growth: An Empirical Investigation for the Federal Republic of Germany

Gebhard Kirchgässner
University of St Gallen

Marcel Savioz
Swiss National Bank and
University of St Gallen

Abstract. *Using quarterly data for the Federal Republic of Germany, we generate four-quarter-ahead forecasts for real GDP growth. Throughout the 1970s and 1980s, other monetary indicators like real M1 or short-run interest rates clearly outperform forecasts which are based on interest rate spreads. This holds for within as well as for ex-post predictions. The same holds for the development after 1992. Moreover, it is shown that simple forecasts based on M1 or on short-run interest rates outperform the common biannual GNP forecasts of the group of German economic research institutes.*

1. INTRODUCTION

Since the 1970s, the German Bundesbank – as well as the Swiss National Bank – has followed a policy oriented toward managing the quantity of money. A precondition for such a policy to be successful is a stable relationship between money and income, as stated by the quantity theory of money. The theoretical foundation of this policy dates back to Friedman (1948), with empirical evidence for the US presented especially in Friedman and Schwartz (1963). Although it took a long time before at least some central banks adopted such a policy, the German and the Swiss attempts were successful.¹ Both national

1. At the beginning of the year 2000, the Swiss National Bank restated its monetary concept adopting inflation forecasts as the main indicator. The importance of monetary aggregates to form inflation forecasts is underlined. In particular, the Swiss National Bank emphasizes the special role of M3 as an important indicator for future inflation. This concept is not identical to inflation targeting in the usual sense, because there is no explicit obligation to achieve a specific inflation target period by period. For a presentation of the Swiss National Bank's monetary policy concept by its former chairman see Meyer (2000).

banks had/have very high reputations for successfully fighting inflation, which was/is apparent in low long-run interest rates in these two countries compared to those of other countries.

However, there is a new literature which cast doubts upon the stability of the relationship between money and (nominal and/or real) income. In a very careful study of the US, Friedman and Kuttner (1992) show that while this relationship holds through the postwar period, it weakens during the 1970s and seems to have collapsed during the 1980s. On the other hand, for the past two decades the spread between the commercial paper rate and the Treasury bill rate seems to have strong predictive power for future movements in real income.² As several other papers also show, interest rate spreads, especially those between long- and short-run rates, seem to become key variables to predict future developments of income and/or prices.³ More recently, the ability of the term structure to predict recessions has been demonstrated.⁴

The empirical evidence presented so far is not restricted to the US; there are also several papers discussing the evidence from Germany. Krämer and Langfeld (1993) show that the interest rate spread between the long- and short-run rates is about four quarters ahead of the change in real GDP, thus serving as a reliable leading indicator.⁵ This is supported by Langfeld (1994), who also points to the predictive power of the term structure for future inflation. Davis and Henry (1994) show that when the yield curve (as well as other spread variables) is used in vector autoregressions, it provides significant information about future macroeconomic activity even after allowing for the effects of other macroeconomic variables. Sauer and Scheide (1995) present evidence that the interest rate spread is Granger-causal to real GDP growth even taking into account the impact of the quantity of money. Gerlach (1997) shows that the interest rate spread contains considerable information about future changes in inflation. And finally, for the period from 1983 to 1989, Harvey (1991) finds that term structure based forecasts of annual economic growth over a horizon of one year outperform the forecasts of the five leading German institutes for economic research.

The evidence for Germany (before its unification in 1990) with respect to the real development is – at least at first glance – rather impressively illustrated in Figure 1, which shows the growth rate of real GDP and the spread between long- and short-run interest rates lagged four periods.⁶ Both curves fit together

2. However, this did not hold for the 1991/92 recession. See Friedman and Kuttner (1998).
3. For the real development see, e.g., Estrella and Hardouvelis (1991), Fayolle and Mathis (1994), Bomfim (1997) and Dotsey (1998); with respect to inflation see Mishkin (1990a, 1990b), Jorion and Mishkin (1991) or Robertson (1992). See also Stambaugh (1988) and Fama (1990).
4. See Estrella and Hardouvelis (1991), Bernard and Gerlach (1998) or Estrella and Mishkin (1998).
5. See also Hesse and Roth (1992) and Ragnitz (1994).
6. The correlation between the two variables is 0.590. The growth rate is calculated as the fourth difference of the logarithms of the levels. Similar figures are presented in Krämer and Langfeld (1993, p. 35) and in Langfeld (1994, p. 3).

Monetary Policy and Forecasts for Real GDP Growth in the FRG

Figure 1 Growth of real GDP and the interest rate spread in the Federal Republic of Germany, 1970–89 (in per cent)

quite well. This especially holds for the two big recessions in 1975 and 1981, whereas in the late 1980s the relationship seems to become weaker.

Accepting this evidence, Gebauer (1989, 1994) recommended that the Deutsche Bundesbank should switch from its quantity-oriented policy to a policy directly influencing the term structure. However, given the record of its policy and the uncertainties connected with a radical policy change, this proposal came at least somewhat surprisingly. Nevertheless, it caused the leading economist of the bank to write a detailed defence of the quantity orientation.⁷

While, with the introduction of the euro, the discussion has lost its relevance for German monetary policy, this issue might be taken up again with respect to the policy of the European Central Bank (ECB), despite the fact that the latter is following a somewhat different 'two-pillared' strategy.⁸ But before a policy which intends to influence the interest rate spread is implemented, three separate questions should be answered positively:

- (i) Is the interest rate spread really a good predictor of real economic activity, and a better one than, for example, the change in the real quantity of money? This presupposes that the relationship between the term structure and future income has a similar lead and is more stable over time than the one between money and income, however stable or unstable the latter

7. See Issing (1994), but also Hesse and Roth (1992). Issing and Hesse were both members of the 'Zentralbankrat', the Board of the German Bundesbank.

8. For the policy of the ECB see ECB (1999) or Issing (1999, 2000).

might be. Despite the evidence mentioned above, this is still debatable. Contrary to the evidence presented in Friedman and Kuttner (1992), the 'tests' presented in most other papers are not comparative and are not very strong. More importantly, there are only few comparisons made with other monetary indicators.⁹

- (ii) Is the term structure just an indicator of monetary policy, or can the central bank directly manipulate it? The latter would imply that it can manipulate both interest rates independently, i.e. that it has an influence on the long-run rate which is independent from its influence on the short-run rate. This has to hold not only in the short run, but also in the long run. Many would deny this possibility; it is at least highly questionable.¹⁰
- (iii) If conditions (i) and (ii) hold, is the Lucas critique irrelevant for the relation between the term structure and income? There are two aspects. First, has the same spread the same effect on real activity, regardless of whether interest rates are high or low? It certainly has a different effect on nominal activity. And second, is this relationship still stable when the central bank tries to exploit it? Whereas Friedman and Kuttner (1992, p. 491), despite their quite significant results, see this as doubtful and, therefore, are not prepared to give strong policy recommendations, the proposal by Gebauer (1994) has to assume such a stability if it is to make sense.

In this paper, using data for the Federal Republic of Germany from 1970 to 1999 we are only concerned with the first question. Taking standard economic wisdom and empirical evidence collected so far, this is the question which is most likely to be answered positively out of all three. However, if it is to be negated, detailed answers to the second and third questions become redundant, at least with respect to monetary policy. Moreover, if this question is denied for Germany, then there is another good reason (besides all the others already mentioned in the literature) that the ECB should follow the example of the Deutsche Bundesbank in its monetary policy (or at least not try to directly influence the term structure).

Like Friedman and Kuttner (1992) did for the US, we look at the performance of a variety of monetary indicators, and not only the term structure, as most of the papers about Germany did.¹¹ Moreover, we explicitly take into account the time-series structure of the data. However, the main difference to the earlier papers about Germany is that we evaluate the performance of the different

9. See, e.g., Davis and Henry (1994) or Ragnitz (1994).

10. Hassler and Nautz (1998) who investigate this question come to the conclusion that such a policy 'has not much chance of success' (p. S.214). For the indicator properties and their role in the future policy of the European Central Bank see also Estrella and Mishkin (1997a).

11. Sauer and Scheide (1995) also use monetary aggregates besides the interest rate spread, but by using all (up to three) lags of the explanatory variables they are unable to conclude whether these variables can be used as predictors over a horizon of four quarters or more. However, as is shown below, the optimal time horizon of forecasts is four quarters for most monetary aggregates. Thus, their results are hardly relevant for our discussion.

indicators by comparing out-of-sample forecasts and not only (within-sample) goodness-of-fit tests.¹² On theoretical grounds, Granger (1980) has pointed out that there can be remarkable differences between the results of within-sample and out-of-sample tests; and in a study about German economic forecasts Kirchgässner (1984) has shown that the evaluation of such forecasts is strongly dependent upon whether one uses within- or out-of-sample tests: rejection of efficiency which seems to be supported by the results of goodness-of-fit tests is no longer possible once we use out-of-sample tests.

The variable to be forecasted is the most prominent indicator of real economic activity, the growth rate of real GDP. Monetary indicators are the real monetary aggregates M1, M2 and M3, long- and short-run interest rates, and two spreads, between interest rates for public and private bonds and between the long- and short-run interest rates. The results for M3 might be especially interesting because the Bundesbank has used M3 as its intermediate target since 1988.

In Section 2, the data and their time-series properties are described. We perform tests for unit roots, cointegration and Granger causality. Section 3 shows the results of goodness-of-fit tests for the various monetary indicators, while in Section 4 the results of out-of-sample tests are presented. These results are in contrast to most previous results regarding Germany: the interest rate spread is not an exceptionally reliable leading indicator for real economic development. This also holds for the period after the unification. Section 5 presents comparisons with the predictions which the group of five (six) German economic research institutes regularly produces every spring and autumn. It shows that these forecasts could have been improved considerably by taking into account the effects of monetary policy as measured by several indicators, but not by the spread. Considering these results, in Section 6 we conclude that there might be some reasons for the institutes to revise their forecast procedures, but not for criticizing the Bundesbank because of its quantity-oriented monetary policy.

2. THE DATA

For real GDP and for the monetary aggregates we use seasonally unadjusted quarterly data from 1970 to 1999.¹³ To get real monetary aggregates we used the (implicit) GDP deflator. For all these variables, we employ logarithmic values ($\ln(\cdot)$), and we use the growth rates compared to the same quarter of the preceding year (Δ_4). From a statistical point of view, if the data are not

12. Out-of-sample forecasts for the Federal Republic of Germany are also presented by Harvey (1991), but he uses only seven (!) observations. For a critique see Hagen and Kirchgässner (1996). Some additional out-of-sample forecasts but only with spread variables and only for the period from 1986 to 1992 are evaluated by Davis and Fagan (1997). They find poor out-of-sample forecasting performance. However, one reason for this might be that this period includes the structural break of the German unification.
13. Description and sources of the data are given in the Appendix.

seasonally integrated one might prefer to use growth rates from quarter to quarter and to model the seasonal structure explicitly. However, in Germany the politically relevant variables which are used in the public discussions (and are also published in the official statistics) are always the growth rates in relation to the same period one year earlier. Thus, despite possible statistical difficulties we have to use these growth rates, if we intend to derive politically relevant results. For the short-run interest rate, we use the three-month money market rate (*GSR*) and the daily rate (*GDR*) in Frankfurt, for the long-run rate the current yield ('Umlaufrendite', *GLR*) of long-term bonds. Because there were regime shifts at the beginning of the 1970s with the abolition of the fixed exchange rate with the US dollar and in 1990 with German unification, we use only the period from 1970 to 1989 to investigate the time-series properties.

To perform the same tests as Friedman and Kuttner (1992) did for the US, we need interest rates of the same maturity for assets issued by private and government agencies. Contrary to the US, however, for Germany such data are not available for money market rates. Therefore, in addition to the weighted interest rate of all long-term bonds (*GLR*) we use long-term rates for publicly (*GLR_o*) and privately (*GLR_p*) issued bonds.

The results of augmented Dickey–Fuller tests for stationarity are given in Table 1.¹⁴ The lag length has been determined by the AIC2 criterion proposed by Pantula *et al.* (1994), i.e. we looked for the minimum of the Akaike criterion and added two lags. Both test procedures lead essentially to the same results. The growth rates of real GDP and of all monetary aggregates are stationary. The same holds for the two short-run interest rates (at least at the 10 percent level) and the interest rate spreads (*GLR*–*GSR*, *GLR*–*GDR*, and *GLR_o*–*GLR_p*). If these variables are stationary, despite the lacking significance the long-run interest rates (*GLR*, *GLR_o*, *GLR_p*) also have to be stationary. Thus, we assume all interest rates to be stationary.

If all growth rates are *I*(0), the levels of real GDP and real money which are *I*(1) might be cointegrated. To check for this possibility we performed Johansen cointegration tests for (the logarithms of) real GDP versus the three real monetary aggregates. We always used five lags. The results for the first eigenvalues are given in Table 2. For the maximum eigenvalue as well as the trace statistic, the null hypothesis of no cointegrating relation can be rejected for M1 and M3 at the 5 per cent level, but for M2 at no conventional significance level.¹⁵ The estimated coefficients of the long-run relations differ,

14. All calculations except those for the cointegration tests have been performed by TSP, Version 4.3, the cointegration tests by PcFiml 8.0. If we employ the Phillips–Perron *z*-statistic instead we get the same results for the significances.

15. The same holds if we apply the small sample correction proposed by Reimers (1992). Using the Schwarz as well as the Hannan–Quinn criterion, the optimal lag length was always *k* = 5. The results are quite insensitive to the number of lags used. If we use two or nine lags instead, we get results which indicate cointegration between real GDP and real M1 and M3 at a significance level of at least 5 per cent. Moreover, if we include seasonal dummies, we still find significant cointegration at least at the 10 per cent level.

Monetary Policy and Forecasts for Real GDP Growth in the FRG

Table 1 Tests for unit roots (1/1970–4/1989, 80 observations)

Variable	<i>k</i>	Dickey–Fuller test	
		<i>r</i>	<i>t</i>
$\Delta_4 \ln(GDP_r)$	5	0.673	–3.428*
$GLR - GSR$	3	0.807	–3.298*
$GLR - GDR$	5	0.688	–3.623**
$GLR_o - GLR_p$	3	0.742	–2.995*
GLR	3	0.906	–2.437
GLR_o	3	0.903	–2.508
GLR_p	3	0.900	–2.432
GSR	3	0.872	–2.845(*)
GDR	4	0.831	–2.988*
$\Delta_4 \ln(M1_r)$	4	0.818	–3.043*
$\Delta_4 \ln(M2_r)$	6	0.791	–3.020*
$\Delta_4 \ln(M3_r)$	9	0.636	–3.345*

Notes: *k* is the maximum lag, *r* the estimated coefficient of the lagged level and *t* the value of the *t*-statistic for $H_0: r = 1$. (*), * or ** shows that the null hypothesis of a unit root can be rejected at the 10, 5 or 1 per cent significance level, respectively.

Table 2 Johansen tests for cointegration (1/1970–4/1989, 80 observations, *k* = 5)

<i>Y</i>	<i>X</i>	Test statistics			
		Eigenvalue	Maximum eigenvalue	Trace	β
$\ln(GDP_r)$	$\ln(M1_r)$	0.191	16.97*	19.12*	0.633
$\ln(GDP_r)$	$\ln(M2_r)$	0.150	13.03	13.43	0.581
$\ln(GDP_r)$	$\ln(M3_r)$	0.202	18.09*	19.20*	0.577

Notes: * shows that the null hypothesis of no cointegration can be rejected at the 5 per cent significance level. β is the coefficient of *X* in the (first) cointegrating equation normalized for *Y*.

however, in all cases quite considerably from 1.0, the value which is to be expected on the basis of the quantity theory. Whatever the reason for this difference between theoretically expected and empirically estimated values might be, it does not necessarily have a negative impact on the forecast properties of the growth rate of real money, while the significant results of the cointegration tests clearly recommend that lagged levels should be included in the forecasting equations.

If a variable is to be of any use for making predictions about future GDP, there should be a simple Granger causal relation between this variable and real GDP. Therefore, we finally tested for Granger causality, employing the direct

Table 3 Tests for Granger causality (1/1970–4/1989, 80 observations, $k = 4$)

Y	X	$F(X \rightarrow Y)$	$F(X \leftarrow Y)$	$F(X - Y)$
$\Delta_4 \ln(GDP_t)$	$GLR - GSR$	2.938*	2.344(*)	0.131
$\Delta_4 \ln(GDP_t)$	$GLR - GDR$	2.560*	1.404	0.175
$\Delta_4 \ln(GDP_t)$	$GLR_o - GLR_p$	0.320	1.098	0.943
$\Delta_4 \ln(GDP_t)$	GLR	2.701*	0.714	0.017
$\Delta_4 \ln(GDP_t)$	GLR_o	2.799*	0.752	0.068
$\Delta_4 \ln(GDP_t)$	GLR_p	3.039*	0.689	0.424
$\Delta_4 \ln(GDP_t)$	GSR	3.960**	1.607	0.246
$\Delta_4 \ln(GDP_t)$	GDR	3.864**	1.108	0.284
$\Delta_4 \ln(GDP_t)$	$\Delta_4 \ln(M1_t)$	8.247***	1.664	0.002
$\Delta_4 \ln(GDP_t)$	$\Delta_4 \ln(M2_t)$	2.676*	1.010	2.400
$\Delta_4 \ln(GDP_t)$	$\Delta_4 \ln(M3_t)$	4.155**	0.365	1.053

Notes: (*), *, ** or *** show that the null hypothesis of non-causality can be rejected at the 10, 5, 1, or 0.1 per cent significance level, respectively.

Granger procedure proposed by Sargent (1976), and we always used four lags. We performed F -tests for simple causal relations from the monetary indicators to real GDP ($F(X \rightarrow Y)$), for reverse causation ($F(X \leftarrow Y)$), and for instantaneous relations between the variables considered ($F(X - Y)$). As (the logarithms of) real GDP and M1 or M3, respectively, are cointegrated, in performing Granger causality tests we included the four periods lagged error correction terms between money and income into the test equations. To calculate these terms, we used the coefficients of the maximum likelihood procedure given in Table 2.¹⁶

According to the results given in Table 3, the spread between the interest rates of public and private bonds is independent of real GDP growth, while the term structure, between a long and a short rate, is Granger causal to it. However, even the latter never has a special predictive value: it is clearly outperformed by the short-run interest rates and by real M1 and M3. Thus, these variables seem to be a better predictor of real GDP than the interest rate spreads. Reverse causation does not seem to exist, which seems to imply, perhaps somewhat surprisingly, that real economic development has no (significant) influence on the Bundesbank policy, that it is not an argument in its reaction function, respectively. It might be, however, consistent with a policy only directed towards price stability. Finally, instantaneous relations do not seem to exist either.

16. As the dependent variable for the Granger causality tests are, in these cases, differences of fourth-order, the corresponding levels have to be fourth-order moving averages, i.e. we first have to transform the levels with the filter $(1 + L + L^2 + L^3)$.

3. GOODNESS-OF-FIT TESTS

To further investigate the predictive quality of the different indicators, we estimated the following equation:

$$Y_t = a_0 + a_1 Y_{t-k} + \sum_{i=1}^4 \beta_i X_{i,t-k} + u_t, \quad k = 1, \dots, 10 \quad (1)$$

with Y being real GDP growth, to find the optimal prediction horizon k^* . As explanatory variables X_i , $i = 1, \dots, 4$, we not only used the spreads and the different monetary aggregates, but also the individual interest rates and the long- and short-run rates together. Moreover, taking the cointegration results of Table 2 into account, we also included the k -periods lagged (logarithmic) levels of the dependent and the explanatory variables. As long as only one explanatory variable was included, k^* is the lag with the highest value of the t -statistic. In those cases where we used two or three explanatory variables, we chose k^* according to the maximum value of the likelihood ratio test for all two or three variables together. Finally, to check whether there is a structural break between the 1970s and the 1980s we used a Chow test to compare the results for these two subperiods.

As simple preliminary OLS regressions show, the optimal lag to predict real GDP growth is $k^*=4$ for the term structure and for all interest rates separately. This corresponds to the earlier results of Hesse and Roth (1992) and Langfeld (1994). Moreover, because the Bundesbank uses a one-year perspective when determining its monetary policy, this is also the lag which is politically most relevant.¹⁷ Thus, we present the regression results for all indicators with $k = 4$.

As we use growth rates compared to the same quarter of the preceding year, the dependent variable as well as the lagged endogenous variable is a third-order moving average on the underlying quarter-to-quarter changes. Moreover, because we look at four-step-ahead predictions, there is another $MA(3)$ component in the residuals. Thus, the residuals might follow an MA -process up to the order of six. The equations for the optimal predictions can nevertheless be estimated with OLS, despite the fact that the 'true' parameters might be inconsistently estimated because of a correlation between the lagged endogenous variable and the error term. For a prediction to be unbiased and – given the available information – even to be optimal in the sense that the prediction error is minimized it is only necessary that the estimated residuals are orthogonal to the explanatory variables. This is just what OLS ensures. On the other hand, to test whether the inclusion of a certain variable into the forecasting equation significantly reduces the forecast error, in calculating the variances the structure of the error term has to be taken into account. Thus, to

17. See, e.g., Herrmann *et al.* (1994).

derive valid *t*-statistics, we performed OLS estimates for the parameters and GMM estimates for the variances.¹⁸

The results are given in Table 4a. We get by far the best result with the maximum of the adjusted R^2 and the minimum of the Schwarz criterion for real M1, taking into account the cointegration between real M1 and real GDP.¹⁹ With respect to single variables, the daily rate performs best, but the three-month rate and the growth rate of real M1 also perform quite well. The differences between these three variables are rather small. On the other hand, M2 seems to have no predictive value at all. The same holds for the spread between public and private bonds. Thus, we can eliminate these variables from further considerations.

The term structure also seems to work quite well. However, if we include the long-run and short-run rates separately, only the short-run rate is significant, whereas the coefficient of the long-run rate in one case even gets the 'wrong' sign. Therefore, it does not make much sense to include the long-run as well as the short-run rate. If we include the single interest rates separately, all are highly significant, but the short-run rates clearly outperform the long-run rate, and they work considerably better than the spread.²⁰ Thus, the short-run rates and not the spread variables are the best interest rate indicators, but all are still outperformed by real M1, as long as the cointegrating relation is taken into account.²¹

If both the short-run interest rates as well as the growth rate of real money have predictive power for real GDP, it seems natural to combine those indicators to test whether the predictive power can be improved further. We have performed such regressions for the two spreads as well as the three single interest rates on the one side, and the three monetary aggregates, with and without the lagged levels, on the other side. The M2 variables have never been significant when included alongside any of the interest rate variables. Moreover, besides M1 and M3 the two spread variables never proved to be significant. Therefore, in Table 4b we only present the results for M1 and M3 on the one side and the three single interest rates on the other side.

In all but one equation, the interest rate variable is significantly different from zero at least at the 10 per cent level. Moreover, measured by the adjusted correlation coefficient as well as the Schwarz criterion the quality of the equations is in all cases improved. The 'best' equation is the one where the long-run rate is included together with the growth rate of M1, taking cointegration into account.²² Thus, the long-run interest rate (and to a lesser

18. This corresponds to the correction proposed by Newey and West (1987). See the description of the GMM procedure in TSP.

19. For the *t*-tests of the coefficients of the lagged levels we used the critical values given by Banerjee *et al.* (1998, p. 276).

20. The same result is reported in Ragnitz (1994).

21. This is in line with results derived by Davis and Fagan (1995) for the European Union.

22. This holds despite the fact that – using the critical values given by Banerjee *et al.* (1998) – the coefficients of the lagged levels are not significantly different from zero.

Monetary Policy and Forecasts for Real GDP Growth in the FRG

Table 4a Forecast regressions for real GDP growth (1/1970–4/1989, 80 observations, $k = 4$)

Explanatory variables X_1, X_2, X_3	β_1	β_2	β_3	$\bar{R}^2$	SC	Chow
$GLR - GSR$	0.695*** (4.74)			0.372	1.214	4.793**
GLR, GSR	0.073 (0.22)	-0.532** (3.32)		0.430	1.157	6.227**
$GLR - GDR$	0.711*** (4.37)			0.335	1.270	3.780*
GLR, GDR	-0.011 (0.04)	-0.563** (3.60)		0.448	1.127	5.870***
$GLR_0 - GLR_p$	0.103 (0.08)			0.021	1.657	2.296(*)
GLR_0, GLR_p	-0.456 (0.40)	-1.364 (1.18)		0.314	1.342	11.081***
GLR	-0.797** (3.20)			0.278	1.352	8.830***
GLR_0	-0.833** (3.35)			0.297	1.326	7.602***
GLR_p	-0.900*** (3.60)			0.320	1.292	10.114***
GSR	-0.503*** (5.10)			0.437	1.103	7.177***
GDR	-0.567*** (5.44)			0.455	1.072	6.270**
$\Delta_4 \ln(M1_r)$	0.331*** (5.47)			0.430	1.116	3.601(*)
$\Delta_4 \ln(M1_r), \ln(GDP_r),$ $\ln(M1_r)$	0.285*** (5.38)	-0.215* (4.26)	0.112 (2.98)	0.605	0.833	1.814
$\Delta_4 \ln(M2_r)$	-0.157 (1.54)			0.085	1.590	1.649
$\Delta_4 \ln(M3_r)$	0.317(*) (1.93)			0.113	1.558	2.113
$\Delta_4 \ln(M3_r), \ln(GDP_r),$ $\ln(M3_r)$	0.304* (2.02)	-0.174(*) (3.45)	0.083 (2.34)	0.227	1.504	2.452*

Notes: The numbers in parentheses are the absolute values of the estimated t -statistics. (*), *, ** or *** denote significance of the estimated coefficients or show that the null hypothesis of no structural break at the end of the 1970s can be rejected at the 10, 5, 1 or 0.1 per cent significance level, respectively. SC is the value of the Schwarz information criterion.

Table 4b Forecast regressions for real GDP growth (1/1970–4/1989, 80 observations, $k = 4$)

Explanatory variables X_1, X_2, X_3, X_4	β_1	β_2	β_3	β_4	$\bar{R}^2$	SC	Chow
$\Delta_4 \ln(M1_r), GLR$	0.273*** (3.61)			-0.284 (1.17)	0.443	1.134	5.164**
$\Delta_4 \ln(M1_r), GSR$	0.173(*) (1.75)			-0.285(*) (1.84)	0.464	1.096	4.889**
$\Delta_4 \ln(M1_r), GDR$	0.148 (1.48)			-0.357* (2.10)	0.470	1.085	4.642**
$\Delta_4 \ln(M1_r), \ln(GDP_r),$ $\ln(M1_r), GLR$	0.210*** (4.16)	-0.141 (2.90)	0.035 (0.91)	-0.709*** (3.71)	0.677	0.673	1.789
$\Delta_4 \ln(M1_r), \ln(GDP_r),$ $\ln(M1_r), GSR$	0.130(*) (1.84)	-0.170(*) (3.47)	0.071 (1.93)	-0.342** (2.94)	0.650	0.752	2.239*
$\Delta_4 \ln(M1_r), \ln(GDP_r),$ $\ln(M1_r), GDR$	0.122 (1.63)	-0.182(*) (3.75)	0.084 (2.32)	-0.360** (2.78)	0.644	0.769	2.520*
$\Delta_4 \ln(M3_r), GLR$	0.144 (0.93)			-0.718** (2.75)	0.286	1.384	6.162***
$\Delta_4 \ln(M3_r), GSR$	0.003 (0.02)			-0.504*** (4.54)	0.430	1.159	6.128***
$\Delta_4 \ln(M3_r), GDR$	0.069 (0.50)			-0.598*** (4.94)	0.451	1.121	6.275***
$\Delta_4 \ln(M3_r), \ln(GDP_r),$ $\ln(M3_r), GLR$	0.098 (0.90)	-0.110 (2.81)	0.002 (0.07)	-1.186*** (5.85)	0.567	0.965	1.161
$\Delta_4 \ln(M3_r), \ln(GDP_r),$ $\ln(M3_r), GSR$	0.015 (0.15)	-0.099 (2.67)	0.014 (0.56)	-0.589*** (7.18)	0.610	0.862	2.133(*)
$\Delta_4 \ln(M3_r), \ln(GDP_r),$ $\ln(M3_r), GDR$	0.076 (0.73)	-0.115 (3.11)	0.030 (1.21)	-0.651*** (7.14)	0.605	0.874	2.212*

Notes: The numbers in parentheses are the absolute values of the estimated t -statistics. (*), *, ** or *** denote significance of the estimated coefficients or show that the null hypothesis of non-structural break at the end of the 1970s can be rejected at the 10, 5, 1 or 0.1 per cent significance level, respectively. SC is the value of the Schwarz information criterion.

extent, the short-run rates as well) seems to contain information which is not included in M1, and it might be possible to exploit this information for predictive purposes. On the other side, the inclusion of M3 does not seem to improve the equations of the interest rates to a large extent: none of the nine estimated coefficients of M3 is significantly different from zero even at the 10 per cent level.

The results of the Chow tests present strong evidence that there was a structural break in the forecasting relations of the interest rate variables, but

Monetary Policy and Forecasts for Real GDP Growth in the FRG

not necessarily of real money. For real M1, the test indicates significance only at the 10 per cent level and, if we include the lagged levels, at no conventional significance level. Thus, there is strong evidence that the relation between interest rates and real growth differs between the 1970s and the 1980s, whereas the relation between money and real growth might have been stable.²³

To get more information about the differences between the two decades we estimated the most important equations for the two decades separately, i.e. for the two spreads, the two short-run rates, and for real M1 with and without the lagged levels. We also estimated these equations for the 1990s. However, the immediate development after unification with its boom in the West and the breakdown of production in the East can hardly be explained by monetary developments. Thus, we can estimate the equations only for the period from 3/1992 onwards, which, until the end of our observation period, 4/1999, leaves us with only 30 observations. But this should be enough to draw some conclusions

The results are given in Table 4c. There is one clear result which holds for all three decades: the spread is always dominated by the single short-run interest rates, and the results for these two variables are always quite close. With respect to all other results there are, however, considerable differences between the three decades. The two spreads have some predictive power in the 1970s, which is increased during the 1980s, but it nearly vanishes in the 1990s. Real M1 has a slightly smaller impact than the single interest rates in the 1970s, it is dominated not only by the short-run rates but also by the two spreads in the 1980s, but it clearly dominates all other variables in the 1990s. If we include the lagged levels, real M1 performs best of all indicators during the 1970s. However, in the 1980s and 1990s the lagged levels no longer prove to be significant. Thus, the influence of the interest rate indicators experienced an inverse U-shaped development over time, their impact increased from the 1970s to the 1980s, but was strongly reduced in the 1990s. This holds in absolute terms as well as compared to the influence of the growth rate of real M1.

The results in Tables 4a to 4c exhibit some clear differences to the results for the US derived by Friedman and Kuttner (1992). First of all, in Germany the spread between public and private bonds has no predictive power at all, whereas in the US the spread between the commercial paper rate and the Treasury bill rate contains significant information about future movements in real GDP. Of course, we employ capital market rates while Friedman and Kuttner (1992) use money market rates, and the maturities differ quite substantially. Nevertheless, if the risk between publicly and privately issued assets differs and if this difference has predictive quality for real economic

23. This seems to contradict the result of Estrella and Mishkin (1997b), who showed that after 1979 in Germany the relation between M3 and real growth vanished. However, this is not really comparable with our result because first we use real and not nominal money growth and second these authors looked only at M3 but not at M1.

Table 4c Forecast regressions for real GDP growth for different subperiods

Explanatory variables	β_1	β_2	β_3	$\bar{R}^2$	SC
1/1970–4/1979, 40 observations					
$GLR - GSR$	0.605** (2.96)			0.289	1.553
$GLR - GDR$	0.614* (2.60)			0.231	1.631
GSR	-0.519*** (3.85)			0.421	1.347
GDR	-0.581*** (3.93)			0.417	1.353
$\Delta_4 \ln(MI_r)$	0.359*** (3.91)			0.406	1.372
$\Delta_4 \ln(MI_r), \ln(GDP_r),$ $\ln(MI_r)$	0.315*** (4.01)	-0.258*** (3.95)	0.120* (2.06)	0.658	0.948
1/1980–4/1989, 40 observations					
$GLR - GSR$	1.150*** (7.31)			0.597	0.457
$GLR - GDR$	1.249*** (6.21)			0.540	0.590
GSR	-0.624*** (8.96)			0.648	0.324
GDR	-0.689*** (9.50)			0.658	0.293
$\Delta_4 \ln(MI_r)$	0.332*** (5.87)			0.487	0.700
$\Delta_4 \ln(MI_r), \ln(GDP_r),$ $\ln(MI_r)$	0.264** (3.07)	-0.066 (0.88)	0.067 (1.14)	0.477	0.847
3/1992–4/1999, 30 observations					
$GLR - GSR$	0.222 (1.35)			0.074	0.653
$GLR - GDR$	0.211 (1.30)			0.071	0.656
GSR	-0.185* (2.14)			0.136	0.584
GDR	-0.182* (2.10)			0.133	0.587
$\Delta_4 \ln(MI_r)$	0.237** (3.87)			0.358	0.287
$\Delta_4 \ln(MI_r), \ln(GDP_r),$ $\ln(MI_r)$	0.224** (2.28)	-0.059 (0.61)	0.016 (0.54)	0.314	0.502

Notes: The numbers in parentheses are the absolute values of the estimated t -statistics. (*), *, ** or *** denote significance of the estimated coefficients or show that the null hypothesis of non-structural break at the end of the 1970s can be rejected at the 10, 5, 1 or 0.1 per cent significance level, respectively. SC is the value of the Schwarz information criterion.

development, this should be effective in both kinds of data. Second, whereas in the US, M1 has lost its predictive power after the 1970s, in Germany it seems to be still maintained, and it improved in the 1990s. And while interest rate variables have gained importance in both countries in the 1980s compared to the 1970s, in Germany the short-run interest rate is by far the most informative interest rate predictor for real GDP, and not the spread, as in the US. Thus, the possibilities for monetary policy seem to have developed quite differently in the US and Germany during this period.

Taking all results together, the goodness-of-fit tests do not indicate a special role for the interest rate spread as four-quarters-ahead predictor of real economic developments in Germany. However, as already mentioned in the Introduction, such within-sample tests are only of limited value in evaluating the forecasting performance of economic indicators. Therefore, in the next section we evaluate out-of-sample forecasts.

4. OUT-OF-SAMPLE FORECASTS

Taking into account the argument by Friedman and Kuttner (1992) for the US, and by Estrella and Mishkin (1997b) for Germany, that there has been a structural break between the 1970s and the 1980s and taking also into consideration the results of the Chow tests presented above, we first concentrated our analysis on the 1980s, i.e. we generated four-quarters-ahead out-of-sample forecasts for the 40 quarters from 1/1980 to 4/1989. To estimate equation (1) we always used 50 observations. For example, we estimated this equation over the period from 4/1966 to 1/1979 and made a (four-quarters-ahead) forecast for 1/1980, re-estimated the equation over the period from 1/1967 to 2/1979 and made a forecast for 2/1980, etc. As indicators, we used the same variables as above, but owing to the poor results we dropped M2 and the spread between the interest rates of public and private bonds, and we did not include the long-run in addition to the short-run interest rate. Moreover, because the differences between the results for the different bond rates reported in Table 4a are small, in the following we only used the general rate (*GLR*). This is also the one which is commonly used in Germany. As a benchmark, we constructed purely autoregressive forecasts where we solely included the four periods lagged value of the dependent variable.²⁴ For these forecasts, we calculated the usual performance measures, the squared correlation coefficient between the actual and predicted values (R^2), the

24. An alternative is to estimate a general autoregressive model and then to construct four-periods-ahead forecasts. To select the optimal lag structure, we estimated such a model over the period from 1970 to 1989. The Akaike criterion as well as the Schwarz criterion showed an optimal lag length of one period (quarter). Thus, for a four-step forecast only the four periods lagged variable is relevant. Thus, one can also directly regress the dependent variable on its four periods lagged value, as we did.

Table 5 Forecast evaluations for real GDP growth (1/1980–4/1989, 40 observations)

Forecasts based on	R^2	$RMSQ$	ME	MAE	β	U
Benchmark (AR)	0.323	2.188	-0.501	1.763	-1.411 (0.61)	0.451
$GLR - GSR$	0.238	1.563	-0.154	1.166	0.834* (0.33)	0.334
$GLR - GDR$	0.212	1.593	0.097*	1.225	0.777 (0.33)	0.356
GLR	0.308	1.902	-1.147	1.504	0.730 (0.26)	0.333
GSR	0.448	1.477	-0.623	1.128	0.828 (0.19)	0.279
GDR	0.459	1.416*	-0.441	1.103*	0.780 (0.17)	0.273*
$\Delta_4 \ln(MI_r)$	0.314	1.589	-0.380	1.227	0.674 (0.21)	0.311
$\Delta_4 \ln(MI_r), \ln(GDP_r),$ $\ln(MI_r)$	0.439	1.553	0.471	1.328	0.641 (0.13)	0.327
$\Delta_4 \ln(MI_r), \ln(GDP_r),$ $\ln(MI_r), GLR$	0.507*	1.764	0.806	1.532	0.568 (0.10)	0.358
$\Delta_4 \ln(MI_r), \ln(GDP_r),$ $\ln(MI_r), GSR$	0.504	1.523	0.616	1.312	0.669 (0.11)	0.324
$\Delta_4 \ln(MI_r), \ln(GDP_r),$ $\ln(MI_r), GDR$	0.497	1.523	0.561	1.323	0.657 (0.11)	0.320

Notes: The numbers in parentheses are the estimated standard errors. * denotes the optimal value with respect to this criterion.

root-mean-squared-error ($RMSE$), the mean error (ME), the mean absolute error (MAE), Theil's inequality coefficient (U) and the regression coefficient of actual on predicted values (β). For this coefficient we also estimated GMM standard errors.

The results are given in Table 5. The forecasts which perform best with respect to a certain criterion are marked by an asterisk (*). For the regression coefficient of actual on predicted values, we evaluate the distance between its value and the theoretically expected value of 1.0. As can be seen, the daily interest rate performs best. It has the smallest $RMSE$, MAE and U , as well as the largest R^2 of all forecasts which use only one variable as indicator. The three-month rate comes very close. With respect to these four indicators, the two short-run rates clearly outperform the two spreads, but also real money. But while real M1 still has some predictive quality, real M3 does not, thus corroborating the results of Estrella and Mishkin (1997b). The combinations of

Monetary Policy and Forecasts for Real GDP Growth in the FRG

Figure 2 $\bar{R}^2$ measures

the two short-run rates with real M1 with or without lagged levels perform with respect to some criteria better, with respect to other criteria worse than the forecasts based on the spreads. With respect to most indicators the performance is considerably worse than those based on the two short-run rates. Thus, there seems to be no rationale to include real money along with interest rates into the forecast equations.

The superiority of the short rate over the term-structure-based forecasts is reflected in the corrected multiple correlation coefficients of the equations employed for these forecasts. Figure 2 shows the development over time of the $\bar{R}^2$ for the three equations using the spread, the (three months) short-run rate and real M1 (including the lagged levels).²⁵ For both forecasts which are based on interest rates we observe an increase in predictive power over time, but the $\bar{R}^2$ of the spread equations is always dominated by the one of the short-run rate. The picture is somewhat different for real M1. Its maximum $\bar{R}^2$ was found in the first half of the 1980s with some decline in their second half. At the end of the 1980s it is outperformed by both interest rate variables.

The same development is reflected in the t -statistics presented in Figure 3.²⁶ Over the whole period, the spread is outperformed by the short-run rate. During most of the 1980s, M1 performs at least better than the spread. This is

25. In Figures 2 and 3, the values for time period t , $t = 1/1980, \dots, 4/1989$, are those of the rolling regressions over the period from $t-54$ to $t-4$ which have been used to construct the prediction for time period t .

26. To make useful comparisons, we plotted the t -value for M1 from the equation without the lagged levels. Thus, we have the same other variables in all three equations compared. According to the results in Table 5, this underestimates the impact of M1.

Figure 3 Absolute values of the t -statistics

contrary to the US experience. However, at the end of the 1980s, M1 is again outperformed by both interest rate variables.

Owing to the structural break and the short period since unification it is impossible to make *ex-post* predictions in the way we did for the period from 1975 to 1989. However, it is possible to estimate equation (1) for the period before unification, e.g., from 1970 to 1989, and to use these estimates and the actual values of the variables at time t to make four-period-ahead forecasts and in this way to predict the development after unification. Excluding as above the transition period from 1990 to mid-1992, we calculated such forecasts for the period from 3/1992 onwards.

The results in Table 6 give strong support for the predictive quality of real money, as soon as we included the lagged levels. With respect to the most important criterion, the root mean squared error, as well as the mean absolute error and Theil's U the M1-based forecasts that account for the cointegrating relation perform best. On the other hand, the two spreads perform rather badly; they do not have any predictive power. The forecasts based only on a single interest rate perform better than those based on a spread, but considerably worse than those based on M1. The R^2 is highest for the combination of real M1 (with cointegration) with the long-run interest rate. However, with respect to the other criteria these forecasts perform considerably worse than the ones based solely on M1. Finally, the predictions based on M3 (with or without cointegration) also perform rather badly.

Thus, the predictive performance of the term structure is even worse for this period than for the time before 1990. On the other hand, the growth rate of M1 together with the lagged levels of M1 and real GDP seems to be the best choice

Monetary Policy and Forecasts for Real GDP Growth in the FRG

Table 6 Forecast evaluations for real GDP growth (3/1992–4/1999, 30 observations)

Forecasts based on	R^2	RMSQ	ME	MAE	β	U
Benchmark (AR)	0.078	1.684	−0.952	1.253	−1.066 (0.80)	0.414
GLR – GSR	0.019	1.663	−0.698	1.411	0.172 (0.27)	0.412
GLR – GDR	0.018	1.574	−0.318	1.300	0.156 (0.26)	0.422
GLR	0.188	2.652	−2.347	2.375	0.530 (0.21)	0.475
GSR	0.128	2.149	−1.648	1.870	0.384 (0.20)	0.433
GDR	0.129	2.057	−1.439	1.809	0.339 (0.17)	0.425
$\Delta_4 \ln(M1_t)$	0.356*	1.920	−1.611	1.621	0.744* (0.20)	0.394
$\Delta_4 \ln(M1_t), \ln(GDP_t),$ $\ln(M1_t)$	0.352	1.378*	0.311*	1.095*	0.459 (0.11)	0.372*
$\Delta_4 \ln(M1_t), \ln(GDP_t),$ $\ln(M1_t), GLR$	0.347	2.014	1.442	1.560	0.433 (0.10)	0.572
$\Delta_4 \ln(M1_t), \ln(GDP_t),$ $\ln(M1_t), GSR$	0.282	1.741	1.036	1.336	0.425 (0.13)	0.513
$\Delta_4 \ln(M1_t), \ln(GDP_t),$ $\ln(M1_t), GDR$	0.265	1.739	0.849	1.317	0.378 (0.12)	0.487

Notes: The numbers in parentheses are the estimated standard errors. * denotes the optimal value with respect to this criterion.

for an equation with predictive quality. This is just the opposite of what we saw in the 1980s, where predictions based on M1 were outperformed by those based on single (short-run) interest rates. In addition, there is nothing in post-1990 experience which would contradict the conclusions from the previous period concerning the comparatively poor forecasting quality of the term structure.

5. COMPARISON WITH FORECASTS OF THE GROUP OF GERMAN ECONOMIC RESEARCH INSTITUTES

The last question to be investigated is the proposition of Harvey (1991) that forecasts based solely on the term structure outperform the common forecasts (*Gemeinschaftsprognosen*) of the group of five major economic research institutes of the Federal Republic of Germany. Since the mid-1960s, this group has

regularly published reports every spring and autumn on the national and international economic developments.²⁷ These reports contain forecasts for the future development of many important macroeconomic variables of the Federal Republic. There are annual forecasts for the current and, in the autumn reports, also for the following year, and semi-annual forecasts for the current, the following and, in the autumn reports, also for the next to the following half-year. For comparison, here we only use the semi-annual predictions for the following half-year to have enough observations to draw serious statistical conclusions, and only the forecast of the growth rate of real GDP (GNP). For these series, we calculated two-period-ahead *ex-post* forecasts, again using rolling regressions.

Even if the forecasts of the institutes are better than those based on the term structure, it might nevertheless be the case that the latter contain information which was not exploited by the forecasts of the institutes. Thus, we will perform tests of efficiency of these forecasts which compare the information contained therein. Moreover, as the term structure forecasts are outperformed by other forecasts based on monetary indicators, we compare the forecasts of the institutes also with the other forecasts constructed above. But, because the forecasts based on combinations of monetary indicators did not perform better than those based on single indicators, we only look at the latter ones.

Up to 1992 the forecasts of the institutes were about GNP and not GDP. Thus, we could not use the forecasts constructed for the evaluations in Table 5 but had to build additional ones for GNP. To do so, we used the monetary variables which were actually available when the forecasts of the institutes were published, e.g. for the forecast of the first half-year which was published in October of the preceding year, we used the interest rates of the third quarter in this year.

When comparing these predictions, one problem arises: which realized values should be used? If we think of our forecasts as predictions over one year, the realized values should be the ones published one year later by the official statistics (which were available for the institutes). However, for constructing the forecasts based on the monetary indicators, we used the corrected (final) values of real GNP which were published, for some periods, only many years later. Thus, whatever series we use, the results are biased in favour of some of the forecast series. To be 'fair' with the institutes, our choice for realized values was to use the data which were available when the institutes made their predictions.

First, we look at the usual evaluation criteria. As the data are now biannual, to have not too few observations we consider the period 1/1975 to 2/1989. The results are given in Table 7. With respect to most criteria the forecasts of the institutes are clearly better than those based on the term structure. However,

27. Members of this group who make the predictions are: Deutsches Institut für Wirtschaftsforschung, Berlin, HWWA-Institut für Wirtschaftsforschung, Hamburg, IFO-Institut für Wirtschaftsforschung, München, Institut für Weltwirtschaft an der Universität Kiel, Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Essen. In 1993, the Institut für Wirtschaftsforschung, Halle, joined this group.

Monetary Policy and Forecasts for Real GDP Growth in the FRG

Table 7 Forecast evaluations for real GNP growth (1/1975–2/1989, 30 observations)

Forecasts based on	R^2	$RMSQ$	ME	MAE	β	U
Institutes	0.277	2.042	−0.303	1.457	0.755 (0.24)	0.344
$GLR - GSR$	0.271	2.125	−0.677	1.459	0.789 (0.26)	0.343
$GLR - GDR$	0.227	2.135	−0.378	1.516	0.701 (0.27)	0.358
GLR	0.330	2.275	−0.993	1.786	0.642 (0.19)	0.337
GSR	0.532	1.796	−0.827	1.362	0.966* (0.17)	0.279
GDR	0.538*	1.706	−0.617	1.297	0.918 (0.17)	0.270*
$\Delta_2 \ln(M1_t)$	0.443	2.043	−0.998	1.389	0.794 (0.18)	0.305
$\Delta_2 \ln(M1_t),$ $\ln(GNP_t), \ln(M1_t)$	0.441	1.929*	−0.355	1.300*	0.673 (0.15)	0.301

Notes: The numbers in parentheses are the estimated standard errors. * denotes the optimal value with respect to this criterion.

while those based on real M3 are clearly worse, the forecasts based on real M1 and on the two short-run interest rates are clearly better than those of the institutes. Thus, the forecasts of the institutes did not seem to incorporate all available information about the monetary policy of the Bundesbank.

The regression coefficients of actual on predicted values and the corresponding standard deviations can be used to perform tests for consistency by regressing the realized on the predicted values.²⁸ The forecasts are consistent and – at least in this respect – weakly rational if the null hypothesis $H_0: \beta = 1$ holds. Moreover, to contain significant information, the alternative null hypothesis $H_0: \beta = 0$ has to be rejected as well. The results presented in Table 7 show that consistency can be rejected neither for the forecasts of the institutes nor for those based on the interest rates; it can only be rejected for those based on M1, if the cointegrating relation is taken into account.²⁹ The alternative hypothesis $H_0: \beta = 0$ has to be rejected for all series except those based on M3.

Tests for efficiency or for strong rationality, respectively, are always restricted to a specific set of information. A seemingly obvious way to do so is to include the relevant variables directly into the regression of the actual on the predicted values and to check whether their coefficients will be significantly different

28. As these are two-step predictions, we again use GMM to calculate the standard errors.

29. For additional tests on the weak rationality of these forecasts see Kirchgässner (1993).

from zero. However, this (often used) approach implicitly assumes that the forecasters had all information over the whole observation period already at the beginning of this period. Thus, this procedure is biased against the forecasts which are tested. The correct way is the recursive construction of forecasts which exploit the relevant information which was available when the actual forecasts (of the institutes) were published and to include both forecasts into the test equation:

$$\Delta_2 \ln(\text{GNP})_{r,t} = \beta_0 + \beta_1 P_{I,t-2}(\Delta_2 \ln(\text{GNP}_{r,t})) + \beta_2 P_{M,t-2}(\Delta_2 \ln(\text{GNP}_{r,t})) + \varepsilon_t \quad (2)$$

In this equation, $P_{I,t-2}$ are the two-step predictions of the institutes and $P_{M,t-2}$ those based on the monetary indicators.

In Kirchgässner (1984), such tests were performed based on the lagged values of real GNP. There the efficiency of the real GNP forecasts of the institutes could not be rejected. Of course, to use the information contained in lagged values of the variable which is to be predicted is an absolutely minimal requirement for efficiency. Thus, the tests performed using the monetary indicators are much stronger.

The results in Table 8 show that the forecasts based on the interest rate spreads do not contain information which is not already used by the institutes: the coefficients of β_2 are not significantly different from zero (and the t -statistics are smaller than those for the institutes' forecasts). Thus, the results of

Table 8 Test for efficiency of real GNP growth forecasts (1/1975–2/1989, 30 observations)

Second forecasts based on	β_0	β_1	β_2	$\bar{R}^2$
<i>GLR – GSR</i>	–0.093 (0.12)	0.436 (1.11)	0.408 (0.97)	0.249
<i>GLR – GDR</i>	0.156 (0.22)	0.591 (1.50)	0.205 (0.50)	0.231
<i>GLR</i>	–0.682 (1.01)	0.524* (2.38)	0.493** (2.82)	0.404
<i>GSR</i>	–0.704 (1.20)	–0.072 (0.26)	1.016*** (3.93)	0.498
<i>GDR</i>	–0.344 (0.62)	–0.099 (0.35)	0.984*** (3.97)	0.506
$\Delta_2 \ln(MI_r)$	–0.391 (0.59)	0.042 (0.13)	0.767** (2.83)	0.402
$\Delta_2 \ln(MI_r),$ $\ln(\text{GNP}_r), \ln(MI_r)$	0.307 (0.51)	0.109 (0.36)	0.616** (2.87)	0.402

Notes: The numbers in parentheses are the absolute values of the estimated t -statistics. (*), *, ** or *** denote significance of the estimated coefficients at the 10, 5, 1 or 0.1 per cent significance level, respectively.

Monetary Policy and Forecasts for Real GDP Growth in the FRG

Harvey (1991) which point in the other direction may be counted as a statistical artefact, caused perhaps by the extremely small number of observations.³⁰ However, the picture changes dramatically if we look at the forecasts based on the two short-run interest rates or on real M1. In all four cases, these forecasts do not only contain information which is not considered in the institutes' forecasts, but the latter forecasts do not contain any information which is not contained in the former ones. Thus, even if the term structure does not work, it remains true that very simple forecasts based on publicly available monetary indicators like the German three-month money market rate would not only have been more precise than the forecasts of the five institutes, but also would have been efficient compared to these forecasts. Even taking into account that the institutes produce forecasts of other macroeconomic variables as well, given the high costs to generate these 'common forecasts' this is (at least) somewhat surprising.

6. SUMMARY AND CONCLUDING REMARKS

In the Federal Republic of Germany, the situation seems to be somewhat different from the US as described by Friedman and Kuttner (1992). Throughout the 1980s, for four-quarter-ahead forecasts of real GDP, short-run interest rates could be used to construct forecasts which were clearly better than those based on interest rate spreads or on real M1. For the period since 1992, however, forecasts based on real M1 outperform not only those based on interest rate spreads but also those based on single interest rates. Thus, while the predictive quality of real M1 might have deteriorated in the 1980s, it seems to have improved again in the 1990s.

Even if we did not find very large differences, the results can differ considerably between within-sample tests of predictive quality and the evaluation of out-of-sample forecasts. Thus, the predictive quality of a model should always be checked using out-of-sample forecasts.

With respect to economic policy, our results provide some additional justification for the quantity-oriented monetary policy which was performed by the Deutsche Bundesbank. Given these results, one might question whether one should stick to M3 or switch to M1 as the main intermediate target of monetary policy. But as long as the main interest of monetary policy is price stability, our results which only look at the relation between real variables do not provide any answer to this question. Moreover, preliminary results show that the price development is generally better predicted using M2 or M3 than using M1.³¹ For a feedback rule concerning real GDP, however, M1 might seem

30. In addition, Harvey (1991), who used annual data, presents some numbers, but not a single test to statistically validate his proposition. For a critical evaluation of the proposition using annual data see Hagen and Kirchgässner (1996).

31. See Kirchgässner and Savioz (1995) or Wolters (1998).

to be the better choice. Of course, this presupposes that the Lucas critique does not hold for this relation, an assumption which is not tested in this paper.

The message for the common forecasts of the five institutes is unambiguous: at least with respect to our observation period from 1975 to 1989 the impact of monetary policy on real GDP has been largely underestimated. This is not that surprising, at least during the 1970s when trust in the effectiveness of fiscal policy was still widespread and monetary aspects played only a minor role in most economic policy discussions. It will be the scope of a future study to check whether the institutes made better forecasts in the 1990s by giving more weight to monetary developments.

APPENDIX: DATA AND SOURCES

Income

GDP Gross domestic product (real)

GNP Gross national product (real)

Source: Deutsches Institut für Wirtschaftsforschung, Berlin, *Vierteljährliche Volkswirtschaftliche Gesamtrechnung*, quarterly data, 1/1960–3/1998.

Money and interest rates

GLR Long-term bond rate (*Umlaufrendite festverzinslicher Wertpapiere*)

GLR_o Long-term rate of government bonds (*Umlaufrendite von Anleihen der öffentlichen Hand*)

GLR_p Long-term rate of privately issued bonds (*Umlaufrendite von Industrieobligationen*)

GSR Three-month money market rate (FIBOR), Frankfurt

GDR Daily rate (*Tagesgeld*), Frankfurt

M1 Quantity of money, M1

M2 Quantity of money, M2

M3 Quantity of money, M3

Source: *Monatsberichte der Deutschen Bundesbank*, different volumes, monthly data, 1/1960–12/1998.

Predictions for GDP and GNP

Source: Arbeitsgemeinschaft deutscher wirtschaftswissenschaftlicher Forschungsinstitute e.V. (ed.), *Die Lage der Weltwirtschaft und der westdeutschen Wirtschaft, Frühjahr 1974–Herbst 1989*, biannual data.

Data are not seasonally adjusted. The quarterly monetary data are averages of the monthly data. The biannual data are averages of the quarterly data. Growth

Monetary Policy and Forecasts for Real GDP Growth in the FRG

rates are calculated as percentage rates compared to the same period of the preceding year.

ACKNOWLEDGEMENTS

A previous version of this paper has been presented at the Konstanz Seminar on Monetary Theory and Monetary Policy. We thank the participants of this discussion, Christian Jochum (UBS, Zürich), as well as two anonymous referees for helpful comments.

REFERENCES

- Banerjee, A., J. J. Dolado and R. Mestre (1998), 'Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework', *Journal of Time Series Analysis* 19, 267–283.
- Bernard, H. and S. Gerlach (1998), 'Does the Term Structure Predict Recessions? The International Evidence', *International Journal of Finance and Economics* 3, 195–215.
- Bomfim, A. N. (1997), 'The Equilibrium Fed Funds Rate and the Indicator Properties of Term-Structure Spreads', *Economic Inquiry* 35, 830–846.
- Davis, E. P. and G. Fagan (1995), 'Indicator Properties of Financial Spreads in the EU: Evidence from Aggregate Union Data', Paper presented at the Tenth Annual Congress of the European Economic Association, Prague, 1–4 September.
- Davis, E. P. and G. Fagan (1997), 'Are Financial Spreads Useful Indicators for Future Inflation and Output Growth in EU Countries?', *Journal of Applied Econometrics* 12, 701–714.
- Davis, E. P. and S. G. B. Henry (1994), 'The Use of Financial Spreads as Indicator Variables: Evidence for the United Kingdom and Germany', *IMF Staff Papers* 41, 517–525.
- Dotsey, M. (1998), 'The Predictive Content of the Interest Rate Term Spread for Future Economic Growth', *Federal Reserve Bank of Richmond Economic Quarterly* 84(3), 31–51.
- ECB (1999), 'The Stability-Oriented Monetary Policy Strategy of the Eurosystem', *ECB Monthly Bulletin* 1, 39–50.
- Estrella, A. and G. A. Hardouvelis (1991), 'The Term Structure as a Predictor of Real Economic Activity', *Journal of Finance* 46, 555–576.
- Estrella, A. and F. S. Mishkin (1997a), 'The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank', *European Economic Review* 41, 1375–1401.
- Estrella, A. and F. S. Mishkin (1997b), 'Is There a Role for Monetary Aggregates in the Conduct of Monetary Policy?', *Journal of Monetary Economics* 40, 279–304.
- Estrella, A. and F. S. Mishkin (1998), 'Predicting U.S. Recessions: Financial Variables as Leading Indicators', *Review of Economics and Statistics* 80, 45–61.
- Fama, E. F. (1990), 'Term-Structure Forecasts of Interest Rates, Inflation, and Real Returns', *Journal of Monetary Economics* 25, 59–76.
- Fayolle, J. and A. Mathis (1994), 'Business Cycles and Interest Rate Spread in the U.S., Europe and Japan: A Stochastic Approach', Document de travail de l'OFCE, No. 94-04, Paris, September.

- Friedman, B. M. and K. N. Kuttner (1992), 'Money, Income, Prices, and Interest Rates', *American Economic Review* 82, 472–492.
- Friedman, B. M. and K. N. Kuttner (1998), 'Indicator Properties of the Paper-Bill Spread: Lessons from Recent Experience', *Review of Economics and Statistics* 80, 34–44.
- Friedman, M. (1948), 'A Monetary and Fiscal Framework for Economic Stability', *American Economic Review* 38, 245–264.
- Friedman, M. and A. J. Schwartz (1963), *A Monetary History of the United States*, Princeton University Press, Princeton, NJ.
- Gebauer, W. (1989), 'Zinsstrukturtheorie und Zentralbankpolitik', in: N. Bub, D. Duwendag and R. Richter (eds.), *Geldwertsicherung und Wirtschaftsstabilität*, Knapp, Frankfurt, pp. 71–86.
- Gebauer, W. (1994), 'Zinsstrukturpolitik', Paper presented at the symposium Zinsstruktur, Geldpolitik und Wirtschaftsentwicklung', Deutsche Bundesbank, Frankfurt, 8 February.
- Gerlach, S. (1997), 'The Information Content of the Term Structure: Evidence for Germany', *Empirical Economics* 22, 161–179.
- Granger, C. W. J. (1980), 'Testing for Causality: A Personal Viewpoint', *Journal of Economic Dynamics and Control* 2, 329–352.
- Hagen, H. M. and G. Kirchgässner (1996), 'Interest Based Forecasts for German Economic Growth', *Weltwirtschaftliches Archiv* 132, 763–773.
- Harvey, C. R. (1991), 'Interest Rate Based Forecasts of German Economic Growth', *Weltwirtschaftliches Archiv* 127, 701–718.
- Hassler, U. and D. Nautz (1998), 'The Link Between German Short- and Long-Term Interest Rates: Some Evidence against a Term Structure Oriented Monetary Policy', *Jahrbücher für Nationalökonomie und Statistik* 217, S.214–226.
- Herrmann, H., H.-E. Reimers and K.-H. Toedter (1994), 'Weighted Monetary Aggregates for Germany', mimeo, Deutsche Bundesbank, Frankfurt, August.
- Hesse, H. and G. Roth (1992), 'Die Zinsstruktur als Indikator der Geldpolitik?', *Kredit und Kapital* 25, 1–25.
- Issing, O. (1994), 'Zinsstruktur oder Geldmenge?', in: H. Sautter (ed.), *Wirtschaftspolitik in offenen Volkswirtschaften*, Vandenhoeck und Ruprecht, Göttingen, pp. 3–21.
- Issing, O. (1999), 'The Eurosystem: Transparent and Accountable or "Willem in Euroland"', *Journal of Common Market Studies* 37, 503–519.
- Issing, O. (2000), 'The Monetary Policy of the European Central Bank: Strategy and Implementation', in: H.-H. Francke, E. Ketzel and H.-H. Kotz (eds.), *Finanzmärkte im Umbruch*, Duncker und Humblot, Berlin, pp. 353–388. (Special Issue 15 of *Kredit und Kapital*.)
- Jorion, P. and F. Mishkin (1991), 'A Multicountry Comparison of Term-Structure Forecasts at Long Horizons', *Journal of Financial Economics* 29, 59–80.
- Kirchgässner, G. (1984), 'Wie gut sind die Prognosen der Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute?', *Weltwirtschaftliches Archiv* 120, 279–300.
- Kirchgässner, G. (1993), 'Testing Weak Rationality of Forecasts with Different Time Horizons', *Journal of Forecasting* 12, 541–558.
- Kirchgässner, G. and M. Savioz (1995), 'Is the Interest Rate Spread a Valid Predictor for Real and Nominal Economic Development? An Empirical Investigation for the Federal Republic of Germany', Discussion Paper No. 9510, University of St Gallen, Department of Economics, October.

Monetary Policy and Forecasts for Real GDP Growth in the FRG

- Krämer, J. W. and E. Langfeld (1993), 'Die Zinsdifferenz als Frühindikator für die westdeutsche Konjunktur', *Die Weltwirtschaft* 1, 34–43.
- Langfeld, E. (1994), 'Die Zinsstruktur als Frühindikator für Konjunktur und Preisentwicklung in Deutschland', Paper presented at the symposium Zinsstruktur, Geldpolitik und Wirtschaftsentwicklung, Deutsche Bundesbank, Frankfurt, 8 February.
- Meyer, H. (2000), 'Zur Geldpolitik im neuen Jahr', mimeo, Swiss National Bank, Zürich.
- Mishkin, F. (1990a), 'The Information in the Longer Maturity Term Structure about Future Inflation', *Quarterly Journal of Economics* 25, 77–95.
- Mishkin, F. (1990b), 'What Does the Term Structure Tell Us About Future Inflation?', *Journal of Monetary Economics* 105, 818–828.
- Newey, W. and K. West (1987), 'A Simple Positive Definite Heteroscedasticity and Autocorrelation Consistent Covariance Matrix', *Econometrica* 53, 703–708.
- Pantula, S. G., G. Gonzales-Farias and W. A. Fuller (1994), 'A Comparison of Unit-Root Criteria', *Journal of Business and Economic Statistics* 12, 449–459.
- Ragnitz, J. (1994), 'Zinsstruktur und Wirtschaftswachstum', *Kredit und Kapital* 27, 11–29.
- Reimers, H.-E. (1992), 'Comparison of Tests for Multivariate Cointegration', *Statistical Papers* 33, 335–359.
- Robertson, D. (1992), 'Term Structure Forecasts of Inflation', *Economic Journal* 102, 1083–1093.
- Sargent, T. J. (1976), 'A Classical Macroeconomic Model for the United States', *Journal of Political Economy* 84, 207–237.
- Sauer, C. H. and J. Scheide (1995), 'Money, Interest Rate Spreads, and Economic Activity', *Weltwirtschaftliches Archiv* 131, 708–722.
- Stambaugh, R. F. (1988), 'The Information in Forward Rates: Implications for Models of the Term Structure', *Journal of Financial Economics* 21, 41–70.
- Wolters J. (1998), 'The Term Structure of Money Growth as a Leading Indicator of Inflation in Germany: An Empirical Analysis for Germany', in: K. H. Oppenländer and G. Poser (eds.), *Social and Structural Change: Consequences for Business Cycle Surveys*, Ashgate, Aldershot, pp. 285–298.