W W P

Wolfsburg Working Papers

Nr. 08 – 04

The Pessimism of Swiss Bond Market Analysts and the Limits of the Sign Accuracy Test

Markus Spiwoks, Nils Bedke und Oliver Hein

Januar 2008

Fachhochschule Braunschweig / Wolfenbüttel
Fachbereich Wirtschaft am Standort Wolfsburg
Robert-Koch-Platz 10-14
D-38440 Wolfsburg
m.spiwoks@fh-wolfsburg.de
nils.bedke@wiwi.uni-goettingen.de
ohein@is-frankfurt.de

The Pessimism of Swiss Bond Market Analysts and the Limits of the Sign Accuracy Test

An empirical investigation of their forecasting success between 1998 and 2007

MARKUS SPIWOKS, NILS BEDKE AND OLIVER HEIN

Corresponding Author:
Prof. Dr. Markus Spiwoks
Wolfsburg University of Applied Sciences
Faculty of Business Administration
Chair in Finance
Robert-Koch-Platz 10-14
D-38440 Wolfsburg
Germany
Phone: +49-5361-83-1511
Fax: +49-5361-83-1502
Mobil: +49-173-65 15 835
m.spiwoks@fh-wolfsburg.de

Ph.D. candidate Nils Bedke
Göttingen University
Faculty of Economic Sciences
Chair in Economic Policy
Platz der Göttinger Sieben 3
D-37073 Göttingen
Germany
nils.bedke@wiwi.uni-goettingen.de

Ph.D. candidate Oliver Hein
Frankfurt University
Faculty of Business Administration
Chair in Information Systems
D-60054 Frankfurt
Germany
ohein@is-frankfurt.de

The Pessimism of Swiss Bond Market Analysts and the Limits of the Sign Accuracy Test

An empirical investigation of their forecasting success between 1998 and 2007

JEL Codes: E47, G12, G21

Keywords: Financial market forecasts, Forecast accuracy, Swiss banks, Swiss economic research institutions

1. Introduction

Forecasting future interest rate developments is of considerable significance for almost all sectors of the banking industry. The high practical relevance of the subject has led to a whole series of studies being carried out which deal with the evaluation of published interest rate forecasts.[1]

Interest rate forecasts which relate to the Swiss money and capital market have only been analyzed by Gosnell and Kolb (1997) until now. However, they only investigate the summarized Consensus Forecasts, in which different forecasts of the individual participating banks are leveled out. Among the ten participating banks there is only a single Swiss bank. In addition, the observation period is rather short (1990-1992). The study therefore leaves many questions unanswered with regard to the forecasting competence of Swiss Bond Market Analysts.

The aim of this investigation is to provide a comprehensive picture of the reliability of Swiss interest rate forecasts. The subjects of analysis are forecasts for 3-month Swiss Franc interest

[1] The most intensive studies have been carried out upon US interest rate forecasts: by Friedman (1980), Throop (1981), Belongia (1987), Dua (1988), Hafer and Hein (1989), Francis (1991), Hafer, Hein, and MacDonald (1992), Domian (1992), Kolb and Stekler (1996), Baghestani, Woo and Zuchegno (2000), Greer (2003), Brooks and Gray (2004), Mose (2005) and Baghestani (2005). British interest rate forecasts have been examined by Scheier and Spiwoks (2006). German interest rate forecasts have been investigated by Albrecht (2000), Spiwoks (2003), Mose (2005) and Benke (2006). Gosnell and Kolb (1997) as well as Spiwoks and Hein (2007) have studied interest rate forecasts for the American, Japanese, British, German, French and Italian money and capital markets.

rate and 10-year Swiss Government bond yield which were published in the period between June 1998 and October 2006 in the journal Consensus Forecasts. The forecasts have forecast horizons of four and 13 months.[2] The period of validity of the forecasts extended from September 1998 to January 2007.

Forecast time series of Swiss banks and research institutes were analyzed. These were the United Bank of Switzerland (UBS), Credit Suisse, Pictet, Julius Bär, Vontobel, Zürcher Kantonalbank, Institut Crea de Macroéconomie Appliquée at the University of Lausanne, the Konjunkturforschungsstelle (KOF) of the Eidgenössische Technische Hochschule Zürich, BAK Basel Economics and the St. Galler Zentrum für Zukunftsforschung (ZZ). The time series were evaluated individually and also as consensus forecasts. With a total of 44 analyzed forecast time series consisting of around 3,650 individual forecasts, this is the first comprehensive investigation of Swiss interest rate forecasts.

The methods used to assess the forecasts include the test for unbiasedness, the efficiency test, comparisons with naïve forecasts and with simple ARIMA models as well as the TOTA coefficient. Particular emphasis, however, is placed upon the sign accuracy test. A specific feature of Swiss bond market analysts in the period of observation, their pronounced pessimism, makes the application of the sign accuracy test impossible in many cases.

The procedures used for analysis are explained in the second chapter. The results of the forecasting quality measurements are presented in the third chapter. The fourth chapter is dedicated to the problems involved with the sign accuracy test. A summary of the results of the study can be found in the fifth chapter.

---

[2] Consensus Forecasts distinguishes between two forecast horizons: three and twelve months. In practical terms, however, the forecast horizons are of four and 13 months. This can be clarified by an example: In the Consensus Forecasts Magazine of September 2001, which comes out in the middle of the month, forecasts for the end of December 2001 and for the end of September 2002 are published. The published forecasts were compiled at the beginning of September at the participating institutions. From the beginning of September to the end of December is actually four months, and from the beginning of September of the year in question to the end of September of the following year is actually 13 months.

2. Methods

The test for unbiasedness examines whether the forecasts correspond to the actual events which take place later on. x_t represents the actual event at the moment in time t, $\hat{x}_t$ represents the forecast of this event, and u_t a residual at the moment in time t.

$$x_t = a + b\ \hat{x}_t + u_t \tag{1}$$

If this relationship is created between the forecast data and corresponding actual events, the following picture arises: It can be stated that the forecasts are unbiased if a does not significantly differ from 0 and b does not significantly differ from 1, and in addition if the error term u is not autocorrelated. The former is verified with the aid of the F-test and the latter by using the Durbin-Watson test. All standard errors are calculated applying the Newey and West (1987) estimation procedure that allows for heteroscedasticity in the error terms. This is indispensably when the forecast horizon is larger than the observational frequency (see Hanson and Hodrick, 1980).

The test for efficiency examines whether appropriate consideration has been given to the actual events which can be observed before the issue of a forecast. x_t represents the actual event at the moment in time t, $\hat{x}_t$ represents the forecast of this event, h the forecast horizon and u_t a residual at the moment in time t.

$$x_t - \hat{x}_t = b_0 + \sum_{i=1}^{4} b_i\ x_{t-h-i} + u_t \tag{2}$$

If the available information has been used efficiently, the analysts' forecast errors should not be correlated with the lags. Following the example of Simon (1989), we take the last four actual events into consideration. Whether an existing correlation between the forecast errors and the lag variables can be viewed as significant is determined with the aid of the F-test.

Let us assume that a black box generates a quantifiable event in regular time intervals. We can observe the time series of these events, but we have no insight whatsoever into the processes occurring inside the black box, and how the visible results were generated. Let us also assume that despite our complete ignorance we have to make a forecast on the future tendency of the time series. As we have no information on the genesis of events, both the future increasing and decreasing course of the time series are equally probable. Thus it seems sensible to assume an unchanged situation in the future (naïve forecast). This idea goes back to the French mathematician Pierre Simon Laplace (1814), who introduced it into the literature as the "principle of insufficient reason". Since then the naïve forecast has been judged as the rock-bottom of forecast quality. Even if nothing is known about the forecast subject, the forecast quality of a naïve forecast can be achieved without effort. If a market expert at least roughly understands the processes to be forecast, his forecasts should have a better quality than naïve forecasts.

Henri Theil (1955, 1966, 1971) used this assumption to develop forecast error measures which allow an implicit comparison of a forecast time series with the time series of the respective naïve forecast. In particular, Theil's new inequality coefficient (Theil's U_2) has been generally accepted.

$$U_2 = \frac{\sqrt{\frac{1}{T-h}\sum_{t=h+1}^{T}(P_t - A_t)^2}}{\sqrt{\frac{1}{T-h}\sum_{t=h+1}^{T}(A_t)^2}} \tag{3}$$

$$P_t = \frac{\hat{x}_t - x_{t-h}}{x_{t-h}} \tag{4}$$

$$A_t = \frac{x_t - x_{t-h}}{x_{t-h}} \tag{5}$$

with

t	=	Continuous time index
T	=	Total amount of present forecasts or actually occurred events
x_t	=	Occurred event at point of time t (t from t = 1 to T)
$\hat{x}_t$	=	Present forecast at point of time t (t from t = 1 to T)
h	=	Forecast horizon
x_{t-h}	=	Occurred event at point of time t-h (point of origin of forecast)

For a perfect forecast, $U_2 = 0$. If $U_2 = 1$ the reviewed forecast time series is, on average, as bad as the time series of naïve forecasts. For $U_2 > 1$ the applied forecasting procedure is even worse than naïve forecasting. A forecast time series which is better than the time series of naïve forecasts will result in $U_2 < 1$.

In addition, it is established whether the performance of the analyzed forecast time series goes significantly beyond a simple ARIMA forecast. The appropriateness of the ARIMA models was determined with the aid of the AIC criterion. The ARIMA model for the 3-month Swiss Franc interest rate contains four autoregressive terms, the consideration of the first differences and four moving average terms. The ARIMA model for the 10-year Swiss Government bond yield contains three autoregressive terms, the consideration of the first three differences and three moving average terms.

The modified Diebold-Mariano test for forecast encompassing is applied here to examine whether the analyzed forecast time series have a level of information content which goes significantly beyond a simple ARIMA forecast. The initial premise here is that a forecasted situation y_k is described by two competing forecast models *i* and *j*:

$$\hat{y}_k = (1-\lambda)\,\hat{y}_{i,k} + \lambda\,\hat{y}_{j,k} \qquad (6)$$

where $0 \leq \lambda \leq 1$. If $\lambda = 0$, then the forecasts generated by model *i* are said to encompass the forecasts generated by model *j*, as model *j* does not contribute any useful information – apart from that already contained in model *i* – to the formation of an optimal composite forecast. Harvey, Leybourne and Newbold (1998) develop a statistic to test the null hypothesis that $H_0 : \lambda = 0$ against the alternative that $H_1 : \lambda > 0$. If the null hypothesis is rejected, then the forecasts contain distinct predictive information which is useful in forming the optimal forecast $\hat{y}_k$.

When forecasts are mainly shaped by the current trend of the variable to be forecast, so that the forecasts correspond to a greater extent with actual events at the time when forecasts were issued than with those at their respective point of time of validity, this is labeled as topically orientated trend adjustment behavior of forecasts (TOTA).

The TOTA coefficient can be used to identify this characteristic. To calculate the TOTA coefficient (see Andres and Spiwoks, 1999; Bofinger and Schmidt, 2003), firstly the coefficient of determination of the forecast data and the actual events are calculated (R^2_A). Then the coefficient of determination of the forecast data from the time when forecasts were issued with the actual events is calculated (R^2_B).

$$\text{TOTA coefficient } = \frac{R^2_A}{R^2_B} = \frac{R^2_{\text{forecasts; actual}}}{R^2_{\text{forecasts; actual - } h}} \quad (7)$$

With h: Forecast horizon

If the value of the TOTA coefficient is < 1, a topically orientated trend adjustment must be assumed. In this case the forecast time series reflects the present more strongly than the future.

Sign accuracy is measured by comparing the forecasts with the actual events and then arranging them in a 2x2 contingency table.

Table 1: 2x2 contingency table

	Actual event: interest rates rise	Actual event: interest rates fall	Σ
Forecast: interest rates rise	N_{11}	N_{12}	$N_{1.}$
Forecast: interest rates fall	N_{21}	N_{22}	$N_{2.}$
Σ	$N_{.1}$	$N_{.2}$	N

The forecasts which estimated the direction of development of interest rates correctly (rising or falling) can be found in the main diagonals (N_{11} and N_{22}). The off-diagonals (N_{12} and N_{21}) contain the forecasts which wrongly estimated the direction of the interest rate change. An χ^2 test is now applied to examine whether the distribution frequency of the four fields is significantly different from a random walk forecast (cf. Diebold and Lopez, 1996; Joutz and Stekler, 2000). If this is the case, it is necessary to determine whether the forecasts examined were significantly better or significantly worse than a random walk forecast.

3. Results

Ballocchi (1998) points out that presenting the data in the form of graphs can be very helpful for the evaluation of forecasts. Here, the consensus forecasts with a forecast horizon of 13 months are examined more closely in order to obtain a first impression of the characteristics of the data. The forecast time series (thin black lines) are compared with the actual interest rate trend (bold black lines) and the naïve forecasts (gray lines). These diagrams already reveal a great deal about the statistical characteristics of the forecast time series.

Figure 1: Ten-year Swiss Government bond yield (bold black line), respective consensus forecasts with 13 months forecast horizon (thin black line) and naïve forecasts (gray line). Source of Data: Consensus Forecasts, Data Stream

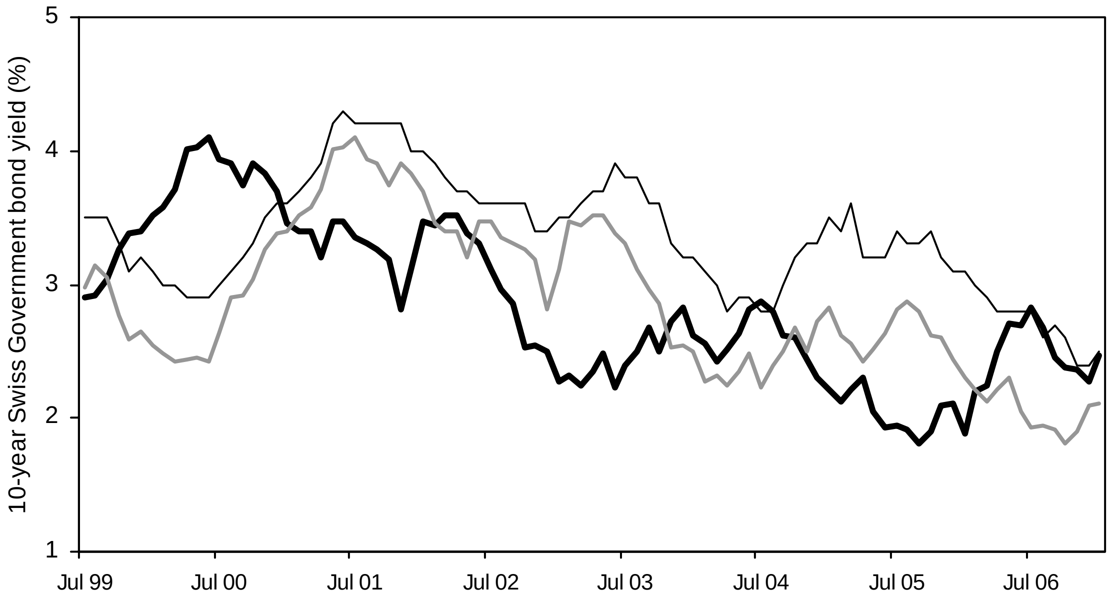


Figure 1 shows that the forecasting efforts with regard to the 10-year Swiss Government bond yield were not very successful overall. Whereas a local interest rate low of 2.9% was forecast for June 2000, in reality there was a high of 4.1%. When the forecast then predicted an interest rate level of 4.2% for November 2001, there was actually a local interest rate low of 2.8%. The forecast predicted a local interest rate high of 3.9% for June 2003. In reality, however, there was a local interest rate low of 2.2%. Whereas the forecast indicated an interest rate level of 3.3% for September 2005, there was actually a low of 1.8%. Overall one can see that the forecast time series hardly corresponds at all to the actual interest rate trends.

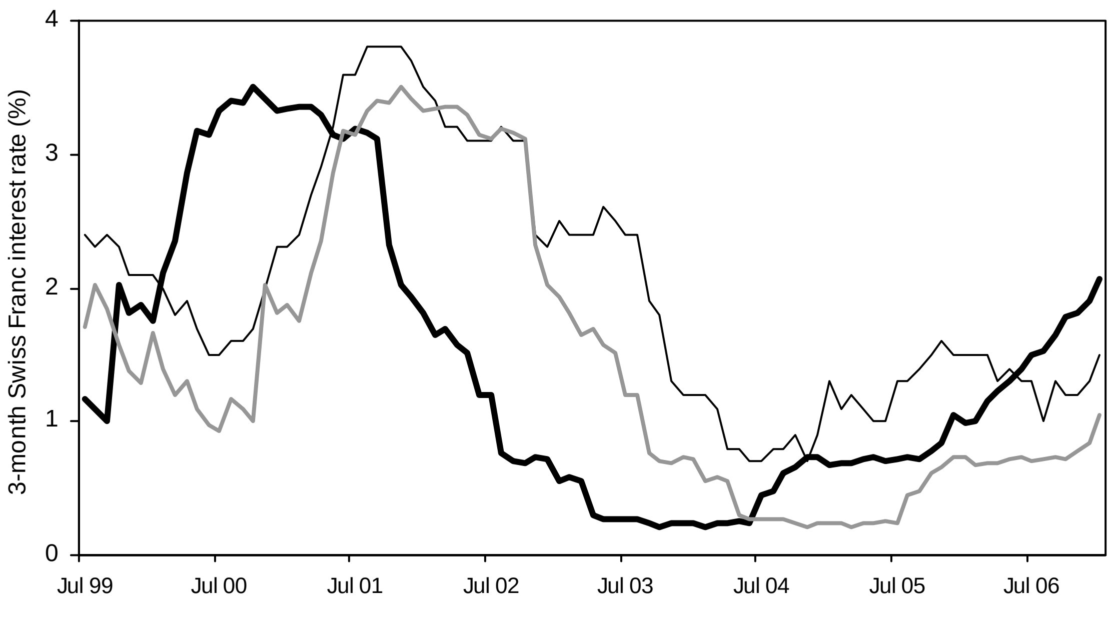
By contrast, the time series of consensus forecasts is very similar to that of the time series of naïve forecasts. This gives rise to the presumption that there is a topically orientated trend adjustment.

It is also noticeable that the curve of the consensus forecasts time series is above that of the time series of naïve forecasts during the entire observation period. Constantly rising interest rates are therefore forecast for the whole period of July 1999 to January 2007. However, the fact that the curve of the time series of naïve forecasts is above the actual interest rate figure for long periods already reveals that interest rates actually fell during more than half of the observation period. The sign accuracy test cannot therefore be expected to indicate successful forecasting work.

It can also be seen that over- and underestimations of the interest rate trend occur over long connected periods. An independent distribution of the residuals is therefore not to be expected. In all probability, the time series of consensus forecasts will not be unbiased.

Analysis of the forecasts for the 3-month Swiss Franc interest rate (Figure 2) reveals similar weaknesses.

Figure 2: Three-month Swiss Franc interest rate (bold black line), respective consensus forecasts with 13 months forecast horizon (thin black line) and naïve forecasts (gray line). Source of Data: Consensus Forecasts, Data Stream



Whereas an interest rate level of 1.6% was forecasted for September 2000, there was actually an interest rate of 3.4%. In October 2002 interest rates were down to 0.7%. The forecast, however, predicated a rate of 3.1% for this period. The forecast predicted a local interest rate high of 2.6% for May 2003. In actual fact, interest rates fell to 0.3%.

Their estimation of the general interest rate trend is not even roughly correct. The steep rise in interest rates in autumn 1999 is not reflected in the forecasts until those for autumn 2000. The next significant increase in spring 2000 is only re-enacted in the forecast time series in spring 2001. The strong fall in interest rates in autumn 2001 is forecast for autumn 2002. The end of decreasing interest rates comes in spring 2003, but is forecast for spring 2004. Here again, the time series of consensus forecasts corresponds much more closely to the time series of naïve forecasts than to the actual interest rate developments. One therefore has to assume again that there is a topically orientated trend adjustment.

The curve of the consensus forecasts time series is always above that of the time series of naïve forecasts except for the short period from March to October 2002. Rising interest rates were therefore expected almost all of the time. However, the curve of the actual interest rate trend is below the time series of naïve forecasts for long periods. Falling interest rates thus occurred much more often than was forecast.

Here again, the under- and overestimations occur over long connected periods of time. It is therefore unlikely that the forecasts can be considered unbiased.

Now that the analysis of the graphs has provided some very clear first impressions, the detailed results are presented below.

Table 2: Results of unbiasedness test of 10-year Swiss Government bond yield forecasts with 13 months forecast horizon

Institution	*a*	st. dev.	*b*	st. dev.	F-dist.	crit. v.	DW	crit. v.
Consensus Forecast	1.938	0.873	0.269	0.238	47.817	3.099	0.083	1.63
BAK Economics	2.397	0.795	0.122	0.202	103.370	3.103	0.078	1.63
Credit Suisse	2.196	0.949	0.194	0.269	35.730	3.111	0.079	1.61
Institut Crea	2.790	0.939	0.029	0.268	25.886	3.130	0.075	1.58
Julius Bär	1.184	0.631	0.438	0.201	87.792	3.138	0.338	1.57
KOF/ETH Zürich	1.736	0.558	0.329	0.146	5.144	3.107	0.098	1.62
Pictet	3.804	1.332	-0.274	0.374	87.395	3.103	0.074	1.63
St. Gallen ZZ	3.126	1.357	-0.015	0.351	33.152	3.153	0.108	1.55
UBS	1.993	0.622	0.252	0.163	61.157	3.100	0.086	1.63
Vontobel	1.009	0.386	0.557	0.091	16.798	3.105	0.131	1.62
Zürcher Kantonalbk.	2.262	0.845	0.190	0.222	43.564	3.111	0.089	1.61

F-test and Durbin-Watson test critical value on 0.05 significance level.

Table 3: Results of unbiasedness test of 10-year Swiss Government bond yield forecasts with 4 months forecast horizon

Institution	*a*	st. dev.	*b*	st. dev.	F-dist.	crit. v.	DW	crit. v.
Consensus Forecast	0.429	0.333	0.798	0.110	13.167	3.089	0.218	1.65
BAK Economics	0.921	0.350	0.627	0.115	20.703	3.092	0.179	1.64
Credit Suisse	0.695	0.353	0.723	0.124	8.340	3.104	0.287	1.62
Institut Crea	0.847	0.460	0.646	0.135	11.959	3.120	0.268	1.64
Julius Bär	0.706	0.272	0.671	0.088	34.567	3.120	0.428	1.60
KOF/ETH Zürich	0.789	0.342	0.665	0.112	19.261	3.095	0.200	1.64
Pictet	0.364	0.335	0.816	0.113	13.996	3.094	0.331	1.64
St. Gallen ZZ	1.028	0.510	0.635	0.151	9.381	3.156	0.214	1.55
UBS	0.663	0.301	0.713	0.095	21.118	3.090	0.318	1.65
Vontobel	0.170	0.338	0.897	0.115	6.534	3.093	0.332	1.64
Zürcher Kantonalbank	0.487	0.273	0.792	0.088	9.851	3.099	0.324	1.63

F-test and Durbin-Watson test critical value on 0.05 significance level.

Table 4: Results of unbiasedness test of 3-month Swiss Franc interest rate forecasts with 13 months forecast horizon

Institution	*a*	st. dev.	*b*	st. dev.	F-dist.	crit. v.	DW	crit. v.
Consensus Forecast	0.551	0.474	0.452	0.215	24.383	3.099	0.047	1.63
BAK Economics	0.844	0.496	0.287	0.225	39.051	3.103	0.044	1.63
Credit Suisse	0.822	0.506	0.323	0.242	19.226	3.115	0.046	1.61
Institut Crea	0.860	0.476	0.031	0.186	32.166	3.340	0.034	1.35
Julius Bär	0.366	0.334	0.364	0.201	74.196	3.138	0.175	1.57
KOF/ETH Zürich	0.769	0.486	0.363	0.216	23.080	3.108	0.043	1.62
Pictet	0.637	0.549	0.386	0.182	36.601	3.103	0.059	1.63
St. Gallen ZZ	0.884	1.262	0.302	0.408	17.216	3.159	0.042	1.55
UBS	0.384	0.423	0.528	0.176	25.974	3.101	0.065	1.63
Vontobel	0.438	0.298	0.628	0.203	8.374	3.103	0.074	1.63
Zürcher Kantonalbank	0.701	0.541	0.419	0.206	21.208	3.109	0.053	1.62

F-test and Durbin-Watson test critical value on 0.05 significance level.

Table 5: Results of unbiasedness test of 3-month Swiss Franc interest rate forecasts with 4 months forecast horizon

Institution	*a*	st. dev.	*b*	st. dev.	F-dist.	crit. v.	DW	crit. v.
Consensus Forecast	-0.022	0.083	0.948	0.057	3.643	3.089	0.269	1.65
BAK Economics	0.034	0.101	0.884	0.077	6.611	3.092	0.213	1.64
Credit Suisse	0.015	0.084	0.928	0.064	3.218	3.103	0.431	1.63
Institut Crea	-0.227	0.173	1.108	0.168	4.439	3.285	0.590	1.40
Julius Bär	-0.039	0.071	0.888	0.053	21.513	3.120	0.821	1.60
KOF/ETH Zürich	0.006	0.127	0.921	0.081	*2.986*	3.097	0.206	1.64
Pictet	0.036	0.095	0.920	0.052	3.693	3.093	0.360	1.64
St. Gallen ZZ	-0.157	0.227	0.966	0.100	4.088	3.156	0.226	1.55
UBS	-0.023	0.085	0.939	0.053	5.633	3.091	0.453	1.65
Vontobel	-0.006	0.066	0.982	0.058	*0.421*	3.093	0.313	1.64
Zürcher Kantonalbank	0.030	0.086	0.945	0.050	*1.685*	3.099	0.417	1.63

F-test and Durbin-Watson test critical value on 0.05 significance level.

Table 6: Results of efficiency test, modified Diebold-Mariano test for forecast encompassing, Theil's U_2 and TOTA coefficient of 10-year Swiss Government bond yield forecasts with 13 months forecast horizon

Institution	Effic. test F-distrib.	Effic. test crit. value	MDM t-distrib.	MDM crit. value	Theil's U_2	TOTA coefficient
Consensus Forecast	3.554	2.482	*1.758*	1.663	1.169	0.046
BAK Economics	4.818	2.486	1.469	1.663	1.513	0.019
Credit Suisse	3.354	2.495	*1.768*	1.664	1.124	0.030
Institut Crea	*1.608*	2.511	*2.281*	1.667	1.180	0.005
Julius Bär	5.335	2.525	1.444	1.668	1.331	0.228
KOF/ETH Zürich	4.428	2.490	*1.757*	1.663	1.234	0.103
Pictet	*0.925*	2.486	*1.715*	1.663	1.466	0.067
St. Gallen ZZ	7.733	2.546	1.457	1.671	*0.965*	0.001
UBS	5.811	2.483	1.577	1.663	1.221	0.064
Vontobel	3.072	2.489	*1.968*	1.663	*0.893*	0.223
Zürcher Kantonalbank	4.012	2.495	*1.902*	1.664	1.174	0.025

Efficiency F-test critical value on 0.05 significance level; Modified Diebold-Mariano test (MDM) on 0.05 significance level.

Table 7: Results of efficiency test, modified Diebold-Mariano test for forecast encompassing, Theil's U_2 and TOTA coefficient of 10-year Swiss Government bond yield forecasts with 4 months forecast horizon

Institution	Effic. test F-distrib.	Effic. test crit. value	MDM t-distrib.	MDM crit. value	Theil's U_2	TOTA coefficient
Consensus Forecast	6.041	2.471	1.271	1.660	1.219	0.561
BAK Economics	5.667	2.474	-0.006	1.661	1.441	0.480
Credit Suisse	3.556	2.487	0.643	1.663	1.218	0.551
Institut Crea	9.576	2.501	*1.890*	1.665	1.785	0.574
Julius Bär	6.304	2.505	0.317	1.665	1.283	0.627
KOF/ETH Zürich	12.450	2.479	*1.759*	1.662	1.572	0.528
Pictet	*1.343*	2.477	1.445	1.661	1.199	0.618
St. Gallen ZZ	13.890	2.560	1.091	1.671	1.282	0.450
UBS	4.344	2.472	1.038	1.661	1.272	0.600
Vontobel	*2.354*	2.475	*2.025*	1.661	1.111	0.610
Zürcher Kantonalbank	*2.354*	2.475	*1.769*	1.662	1.151	0.645

Efficiency F-test critical value on 0.05 significance level; Modified Diebold-Mariano test (MDM) on 0.05 significance level.

Table 8: Results of efficiency test, modified Diebold-Mariano test for forecast encompassing, Theil's U_2 and TOTA coefficient of 3-month Swiss Franc interest rate forecasts with 13 months forecast horizon

Institution	Effic. test F-distrib.	Effic. test crit. value	MDM t-distrib.	MDM crit. value	Theil's U_2	TOTA coefficient
Consensus Forecast	10.780	2.482	0.666	1.662	1.015	0.154
BAK Economics	14.461	2.486	-0.104	1.663	1.221	0.081
Credit Suisse	7.565	2.501	0.334	1.665	1.030	0.080
Institut Crea	7.806	2.759	*3.290*	1.699	2.674	0.009
Julius Bär	13.072	2.525	0.292	1.668	1.239	0.256
KOF/ETH Zürich	12.310	2.492	0.703	1.663	1.276	0.109
Pictet	6.738	2.486	0.579	1.663	1.597	0.168
St. Gallen ZZ	9.058	2.553	-0.001	1.672	1.365	0.043
UBS	10.280	2.484	1.039	1.662	1.026	0.251
Vontobel	12.261	2.486	*2.988*	1.663	*0.688*	0.342
Zürcher Kantonalbank	9.058	2.553	0.950	1.664	1.015	0.161

Efficiency F-test critical value on 0.05 significance level; Modified Diebold-Mariano test (MDM) on 0.05 significance level.

Table 9: Results of efficiency test, modified Diebold-Mariano test for forecast encompassing, Theil's U_2 and TOTA coefficient of 3-month Swiss Franc interest rate forecasts with 4 months forecast horizon

Institution	Effic. test F-distrib.	Effic. test crit. value	MDM t-distrib.	MDM crit. value	Theil's U_2	TOTA coefficient
Consensus Forecast	6.023	2.471	*2.314*	1.660	*0.927*	0.850
BAK Economics	12.331	2.474	-0.519	1.661	1.075	0.783
Credit Suisse	4.080	2.486	*2.029*	1.663	*0.890*	0.854
Institut Crea	5.240	2.690	*1.848*	1.691	2.728	1.061
Julius Bär	2.722	2.505	*2.188*	1.665	1.012	0.952
KOF/ETH Zürich	14.800	2.479	-0.114	1.662	1.257	0.783
Pictet	*1.623*	2.475	*3.206*	1.661	1.227	0.900
St. Gallen ZZ	6.738	2.550	-0.230	1.671	1.398	0.730
UBS	4.237	2.473	*2.378*	1.661	*0.926*	0.902
Vontobel	3.387	2.475	*2.754*	1.661	*0.797*	0.888
Zürcher Kantonalbank	3.660	2.482	*2.738*	1.662	*0.993*	0.883

Efficiency F-test critical value on 0.05 significance level; Modified Diebold-Mariano test (MDM) on 0.05 significance level.

The results of the test for unbiasedness are unambiguous (Tables 2-5). The F-test indicates that a significantly deviates from 0 and/or b significantly deviates from 1 in 41 out of 44 forecast time series. In addition, the Durbin-Watson test shows that the residuals are not independently distributed in any of the 44 time series. None of the 44 time series can therefore be considered unbiased.

At least the efficiency test leads to favorable results in individual cases (Tables 6-9). All of the available information was efficiently taken into account in six of the 44 forecast time series analyzed (13.6%).

The comparison of the forecast time series with the forecast value of a simple ARIMA model (Tables 6-9) provides reasonable results. In just under half of the cases (47.7%) it is revealed that the analyzed forecast time series exhibit an explanatory value which is significantly above that of the ARIMA model.

However, only eight out of 44 forecast time series (18.2%) bear comparison with the naïve forecasts (Tables 6-9). In these cases, Theil's inequality coefficient U_2 is below the threshold value of 1.

The picture provided by the TOTA coefficient is sobering (Tables 6-9). There is a topically orientated trend adjustment in all 44 forecast time series. They therefore reflect the interest rate trend at the time the forecast was made much more strongly than at the time which the forecasts were intended for. The forecasts are considerably closer to the naïve forecasts than to the actual interest rate trend.

4. The limits of the sign accuracy test

The sign accuracy test is a common and sound procedure for assessing the quality of forecasts. The test analyzes whether the future development tendency (rising or falling) of the subject of the forecast is significantly better grasped than by a random walk forecast (see Chapter 2 for more details).

The limits of the sign accuracy test are reached when the forecasts are very one-sided. For example, if bond market analysts are very pessimistic and therefore expect predominantly rising interest rates (= falling bond prices), this can result in such low values in the 2x2 contingency table (with the frequencies expected for random processes) that there is no longer a convergence of the test statistic with the chi-square distribution. To illustrate this, here is the example of the Zürcher Kantonalbank and its 10-year Swiss Government bond yield forecasts with a forecast horizon of 13 months.

Table 10: 2x2 contingency table for the actual frequencies of the Zürcher Kantonalbank's 10-year Swiss Government bond yield forecasts with a forecast horizon of 13 months

	Actual interest rate trend ≥ 0	Actual interest rate trend < 0	Σ
Forecast interest rate trend ≥ 0	39	42	81
Forecast interest rate trend < 0	1	0	1
Σ	40	42	82

The Zürcher Kantonalbank thus only forecast falling interest rates a single time, although in actual fact rates fell in 42 out of 82 cases (see Table 10). This can fairly be described as marked pessimism. Starting out from this situation, we obtain the following contingency table with the frequencies expected for random processes (Table 11).

Table 11: 2x2 contingency table with the frequencies expected for random processes for Zürcher Kantonalbank's 10-year Swiss Government bond yield forecasts with a forecast horizon of 13 months

	Actual interest rate trend ≥ 0	Actual interest rate trend < 0	Σ
Forecast interest rate trend ≥ 0	39.51	41.49	81
Forecast interest rate trend < 0	0.49	0.51	1
Σ	40	42	82

Two fields in the contingency table (with the frequencies expected for random processes) thus exhibit values of < 1. As a consequence of this, the test statistic no longer even approximately corresponds to the chi-square distribution. Some authors assume that all the fields of the contingency table have to exhibit values > 1. Others demand that all the fields should exhibit frequencies of > 5. For this study, this means that the sign accuracy test partly provides no results (in the case of frequencies ≤ 1) and partly provides results which are subject to reservations (in the case of frequencies between 1 and 5).

Table 12: Results of sign accuracy test

	Ten-year Swiss Government bond yield forecasts				Three-month Swiss Franc interest rate forecasts			
	13 month forecast horizon		4 month forecast horizon		13 month forecast horizon		4 month forecast horizon	
Institution	χ^2-dist.	result	χ^2-dist.	result	χ^2-dist.	result	χ^2-dist.	result
Consensus Forec.	n.a.	n.a.	0.278	o	*7.138	+	12.578	+
BAK Economics	n.a.	n.a.	8.472	-	n.a.	n.a.	0.159	o
Credit Suisse	*0.164	o	0.961	o	*4.580	+	8.351	+
Institut Crea	8.279	+	1.686	o	n.a.	n.a.	*3.344	o
Julius Bär	n.a.	n.a.	*0.018	o	12.774	+	20.649	+
KOF/ETH Zürich	*4.504	+	0.145	o	7.534	+	0.004	o
Pictet	n.a.	n.a.	0.270	o	*6.650	+	10.999	+
St. Gallen ZZ	*4.325	+	0.432	o	*0.017	o	*5.146	-
UBS	n.a.	n.a.	1.100	o	*8.261	+	18.249	+
Vontobel	10.968	+	0.126	o	40.842	+	15.271	+
Zürcher Kant.bk.	n.a.	n.a.	0.026	o	*14.548	+	5.780	+

Critical value on 0.05 significance level = 3.8414; o = not significantly different from a random process; + = significantly better than a random process; - = significantly worse than a random process; * = result subject to reservations (in the case of frequencies 1 and 5), n.a. = not available (in the case of frequencies ≤ 1).

Only in 12 out of 44 forecast time series can it be assumed that the future trend (rising or falling interest rates) has been grasped significantly better than by a random walk forecast (Table 12). On balance, the forecasts for the trend of the 3-month Swiss Franc interest rate are more successful than the forecasts for the trend of the 10-year Swiss Government bond yield.

In eight out of 44 forecast time series no result could be achieved because at least one field in the contingency table (with the frequencies expected for random processes) exhibited a value of ≤ 1. This can be traced back to the one-sided expectations of the forecasters, who always thought that interest rates would rise. UBS, for example, issued 90 forecasts for the 10-year Swiss Government bond yield with a forecast horizon of 13 months. In 50 of these 90 cases, interest rates fell in the forecast period. UBS, however, did not make a single forecast predicting falling interest rates. The situation is similar with BAK Basel Economics, which made 88 forecasts. In 49 of these cases, interest rates fell. BAK Basel Economics, however, forecast rising rates in all 88 cases. Zürcher Kantonalbank made 81 forecasts, 80 of them for rising interest rates, although in reality interest rates fell in 42 cases during the forecast period. The only forecast for falling rates was ironically made for a period in which interest rates increased (see Table 10). The situation is largely the same in all eight cases in which the sign accuracy test could not be carried out.

5. Conclusion

Overall it is fair to say that there is a considerable need for improvement in the forecasts of Swiss interest rate trends. This becomes particularly clear when the institutes are ranked according to a scoring procedure and the forecast time series of the best institutes are then analyzed.

A very simple distribution of the scoring points is carried out: for each successfully completed evaluation procedure (1. Test for unbiasedness 2. Efficiency test 3. Comparison with a naïve forecast 4. Comparison with the ARIMA model 5. TOTA coefficient and 6. Sign accuracy test) one point is awarded. An institute can thus obtain a maximum total of 24 points for the four different forecasting subjects.

Table 13: Ranking of the institutes

Ranking	Institute	*Points*	Ranking	Institute	*Points*
1.	Vontobel	*11*	6.	Julius Bär	*3*
2.	Institut Crea	*6*		KOF/ETH Zürich	*3*
	Pictet	*6*		UBS	*3*
	Zürcher Kantonalbank	*6*	9.	St. Gallen ZZ	*1*
5.	Credit Suisse	*4*	10.	BAK Economics	*0*

With eleven out of 24 possible points, Bank Vontobel is the clear winner in this rating (Table 13). Three aspects of Bank Vontobel's results are particularly noteworthy:

1. Three of the four forecast time series predict the future interest rate trend more precisely than the alternative of naïve forecasts.

2. All four forecast time series have an information content which goes significantly beyond that of the ARIMA models.

3. Three of the four forecast time series estimate the future direction of the trend (rising or falling) significantly better that a random walk forecast.

This result is also very respectable in an international comparison. Nevertheless, a glance at the curve of Vontobel's forecast time series reveal that the winner of the scoring procedure also has room for improvement.

Figure 3: Ten-year Swiss Government bond yield (bold black line), respective forecasts of Vontobel with 13 months forecast horizon (thin black line) and naïve forecasts (gray line). Source of Data: Consensus Forecasts, Data Stream

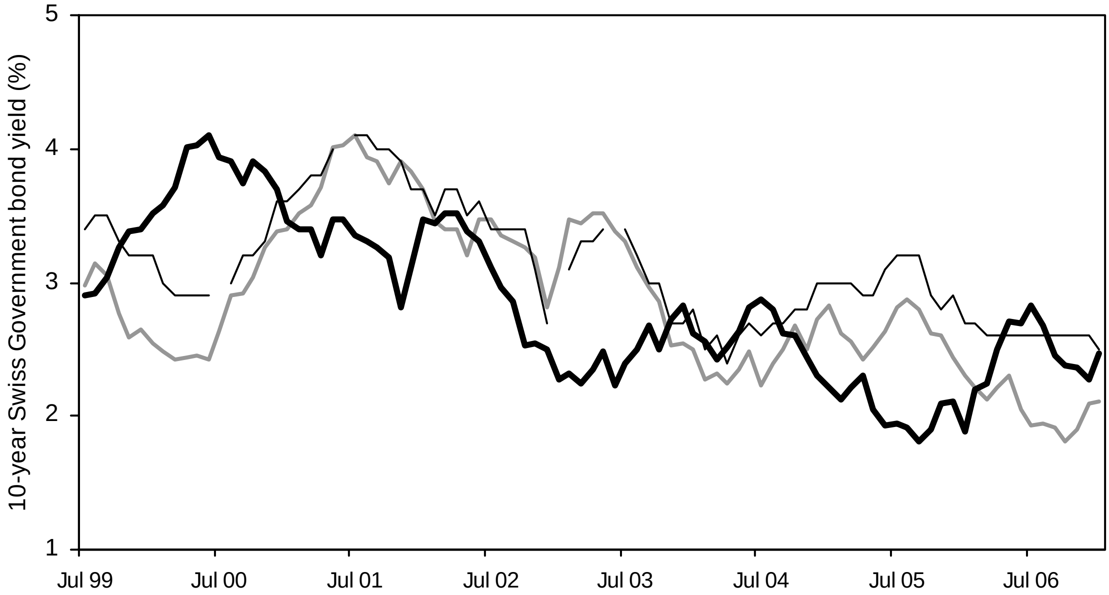
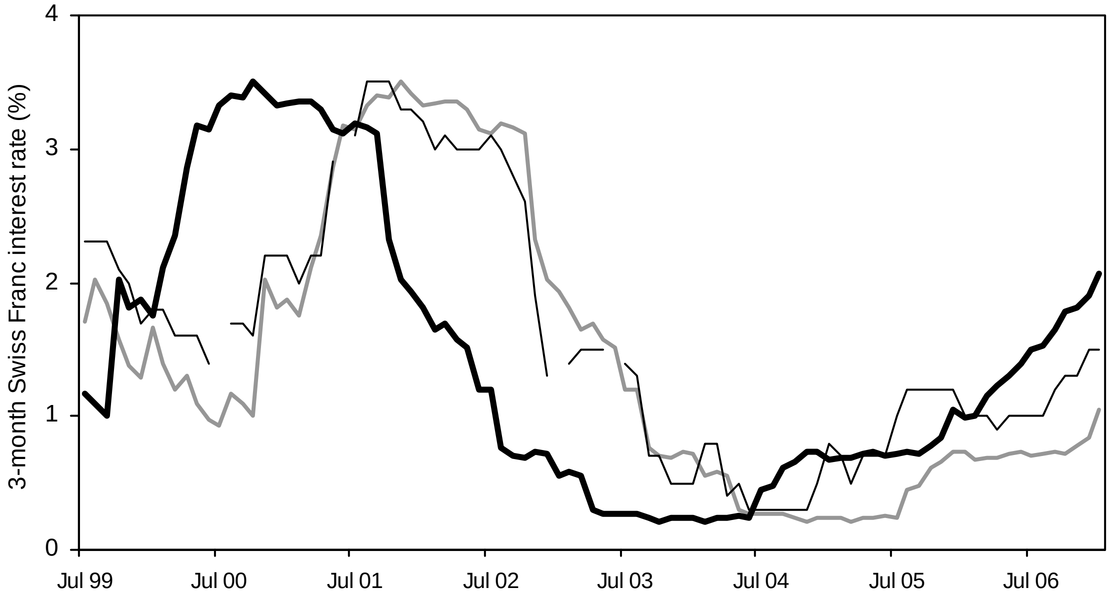


Figure 4: Three-month Swiss Franc interest rate (bold black line), respective forecasts of Vontobel with 13 months forecast horizon (thin black line) and naïve forecasts (gray line). Source of Data: Consensus Forecasts, Data Stream



Both Figure 3 and Figure 4 show that Vontobel's forecast time series correspond much more closely to the naïve forecasts than to the actual interest rate trend. Against this background, it is not surprising that the results of the test for unbiasedness and the results of the TOTA coefficient reveal that there is still a fundamental need for action even for the winner of the scoring procedure.

References

ALBRECHT, T. (2000). *Zur Eignung professioneller Zinsprognosen als Entscheidungsgrundlage*. Diskussionsbeiträge der Sonderforschungsgruppe Institutionenanalyse (sofia), 00-7, Darmstadt.

ANDRES, P. and SPIWOKS, M. (1999). "Forecast Quality Matrix: A Methodological Survey of Judging Forecast Quality of Capital Market Forecasts". *Journal of Economics and Statistics* 219: pp. 513-542.

BAGHESTANI, H. (2005). "Improving the Accuracy of Recent Survey Forecasts of the T-bill Rate". *Business Economics* 40: pp. 36-40.

BAGHESTANI, H., JUNG, W. and ZUCHEGNO, D. (2000). "On the information content of futures market and professional forecasts of interest rates". *Applied Financial Economics* 10: pp. 679-684.

BALLOCCHI, G. (1998). „Die statistische Evaluation von Finanzmarktprognosen". In: Kutscher, C. and Schwarz, G. (eds), *Aktives Portfolio Management, Der Anlageentscheidungsprozeß von der Prognose bis zum Portfolio*, Zürich, pp. 23-49.

BELONGIA, M. T. (1987). "Predicting Interest Rates". *Federal Reserve Bank of St. Louis Review* March: pp. 9-15.

BENKE, H. (2006). "Was leisten Kapitalmarktprognosen? – die Sicht eines Stiftungsmanagers". *Zeitschrift für das gesamt Kreditwesen* 35: pp. 902-906.

BOFINGER, P. and SCHMIDT, R. (2003). "On the Reliability of Professional Exchange Rate Forecasts: An Empirical Analysis for the €/US-$ Rate". *Financial Markets and Portfolio Management* 17: pp. 437-449.

BROOKS, R. and GRAY, B. (2004). "History of the Forecasters – Naive forecasts are better than the consensus WSJ forecast". *Journal of Portfolio Management* 31: pp. 113-117.

DIEBOLD, F. X. and LOPEZ, J. A. (1996). *Forecast Evaluation and Combination*, in: Maddala, G. S. and Rao, C. R. (eds), Handbook of Statistics. 14: pp. 241-268, Amsterdam.

DOMIAN, D. L. (1992). "Money Market Mutual Fund Maturity and Interest Rates". *Journal of Money, Credit, and Banking* 24: pp. 519-527.

DUA, P. (1988). "Multiperiod Forecasts of Interest Rates". *Journal of Business and Economic Statistics* 6: pp. 381-384.

GOSNELL, T. F. and KOLB, R. W. (1997). "Accuracy of International Interest Rate Forecasts". *The Financial Review*. 32: pp. 431-448.

FRANCIS, J. (1991). "Management Anticipation of Interest Rates: The Case of Commercial Banks". *Journal of Business Finance and Accounting* 18: pp. 675-695.

FRIEDMAN, B. M. (1980). "Survey Evidence on the 'Rationality' of Interest Rate Expectations". *Journal of Monetary Economic* 6: pp. 453-465.

GREER, M. (2003). "Directional accuracy tests of long-term interest rate forecasts". *International Journal of Forecasting* 19: pp. 291-298.

HAFER, R. W. and HEIN, S. E. (1989). "Comparing Futures and Survey Forecasts of Near-Term Treasury-Bill Rates". *Federal Reserve Bank of St. Louis Review* May/June: pp. 33-42.

HAFER, R. W., HEIN, S. E. and MACDONALD, S. S. (1992). "Market and Survey Forecasts of the Three-Month Treasury-Bill Rate". *Journal of Business* 65: pp. 123-138.

HANSON, L. P. and HODRICK, R. J. (1980). "Forward Exchange Rates as Optimal Predictors of Future Spot Exchange Rates: An Econometric Analysis". *Journal of Political Economy* 88: pp. 829-853.

HARVEY, D. I., LEYBOURNE, S. J. and NEWBOLD, P. (1998). "Tests for Forecast Encompassing". *Journal of Business & Economic Statistics* 16: pp. 254-259.

JOUTZ, F. and STEKLER, H. O. (2000). "An evaluation of the predictions of the Federal Reserve". *International Journal of Forecasting* 16: pp. 17-38.

KOLB, R. A. and STEKLER, H. O. (1996). "How well do analysts forecast interest rates?" *Journal of Forecasting* 15: pp. 385-394.

LAPLACE, P. S. (1814), *Essai philosophique sur les probabilites*, Paris.

MOSE, J. S. (2005). „Expert Forecasts of Bond Yields and Exchange Rates". *Danmarks Nationalbank Monetary Review* 4th Quarter: pp. 91-95.

NEWEY, W. K. and WEST, K. D. (1987). „A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". *Econometrica* 55: pp. 703-708.

SCHEIER, J. and SPIWOKS, M. (2006). *Aktives Portfoliomanagement am britischen Anleihenmarkt*. Frankfurt.

SPIWOKS, M. (2003). „Qualität der Zinsprognosen deutscher Banken: Eine empirische Analyse". *Kredit und Kapital* 36: pp. 289-308.

SPIWOKS, M. and HEIN, O. (2007). "Die Währungs-, Anleihen- und Aktienmarktprognosen des Zentrums für Europäische Wirtschaftsforschung – Eine empirische Untersuchung des Prognoseerfolges von 1995 bis 2004". *Wirtschafts- und Sozialstatistisches Archiv* 1: pp. 43-52.

THEIL, H. (1955). *Who forecasts best?*, International Economic Papers, No. 5.

THEIL, H. (1966). *Applied Economic Forecasting*, Amsterdam.

THEIL, H. (1971). *Principles of Econometrics*, New York.

THROOP, A. W. (1981). "Interest Rate Forecasts and Market Efficiency". *Economic Review of the Federal Reserve Bank of San Francisco*. Spring: pp. 29-43.