
Interest rate forecasting and the financial crisis: a turning point in more than just one way

Frederik Kunze* and Mario Gruppe

Norddeutsche Landesbank Girozentrale,
Department of Economics,
Friedrichswall 10, 30159 Hannover, Germany
Email: frederik.kunze@nordlb.de
Email: mario.gruppe@nordlb.de
*Corresponding author

Tilo Wendler

Hochschule für Technik und Wirtschaft (HTW) Berlin,
Fachbereich Wirtschaftswissenschaften I,
Treskowallee 8, 10318 Berlin, Germany
Email: Wendler@htw-berlin.de

Abstract: Investors frequently rely on forecasts published by professional analysts. During the financial crisis uncertainty has substantially risen. Not surprisingly experts' predictions should be more than welcome for decision makers. After modelling the long-term relationship between the three month EURIBOR and the Consensus Economic forecast by using co-integration analysis this paper tests for changes in the accuracy of forecasts for the three month EURIBOR. We use traditional evaluation methods like sign accuracy tests, turning point analysis and the root mean square error (RMSE). We find evidence for a crisis related structural break. Checking for quality changes it can be stated that the forecasters' accuracy after the first months of the crisis is not much better in general. Furthermore, neither before nor after the structural break the analysts' forecasts did outperform a random prediction. But it can at least be stated that there is a significant improvement in the turning point prediction.

Keywords: financial crisis; interest rate forecasts; co-integration; structural breaks; forecast evaluation; turning points; TOTA-coefficient.

Reference to this paper should be made as follows: Kunze, F., Gruppe, M. and Wendler, T. (2015) 'Interest rate forecasting and the financial crisis: a turning point in more than just one way', *Int. J. Financial Engineering and Risk Management*, Vol. 2, No. 1, pp.1–16.

Biographical notes: Frederik Kunze is currently working as an Economist in Norddeutsche Landesbank. He studied Economics from the Georg-August-University in Goettingen as well as in Nottingham (UK). His main fields of interest are the Chinese economy, the interaction of macroeconomic variables and financial markets and especially the evaluation of interest rate forecasts.

Mario Gruppe works as an Economist in Norddeutsche Landesbank and Teaches macroeconomics and microeconomics in the VWA Braunschweig. He

studied Economics and Political Sciences from the University of Erfurt. His main fields of interest are the European Monetary system, financial markets and central banks.

Tilo Wendler is a Professor for Quantitative Methods in HTW Berlin; University of Applied Sciences. He holds a Doctoral degree from the University of Klagenfurt. Before his career in higher education he worked for the Association of German Public Sector Banks. His research interests are focused on the field of financial econometrics and data analysis.

1 Introduction

For decades financial market participants (e.g., asset managers or treasurers) focus on the monetary policy makers and the impact of their decisions on the yield curve. Decision makers in the financial industry have regularly to rely on their own forecasts or on predictions published by other financial market professionals. Especially short term interest rates like the three month EURIBOR are said to be directly affected by the ECB's benchmark rate and the changes thereof. Because of that forecasts for the three months EURIBOR should be directly influenced by the expected interest rate moves of the ECB.

In times of financial crisis and rising uncertainty accurate forecasts are needed the most. This raises the question whether the quality of interest rate forecasts in times of financial crisis is sufficient. During the subprime crisis followed by financial turmoil and the European sovereign debt crisis central banks had to find quick answers. With respect to the performance of interest rate forecasts for the three month EURIBOR one might argue that this has to be a reason for a deteriorating performance of the forecasters' predictions. But that has not necessarily to be true. With the central banks – not only in Europe – becoming a paramedic for the global economic activity the lowering of the ECB's benchmark rate might have been broadly expected by analysts with the outbreak of the financial crisis in 2008. In the remainder of this paper we will focus on that aspect using traditional forecasts evaluation methods like sign accuracy tests, turning point analysis and the root mean square error to test for changes in the quality of interest rate forecasts. Within our assessment we put a special emphasis on the accuracy of turning points of the three month EURIBOR. Additionally we test for Topically Orientated Trend Adjustment (TOTA) and perform the commonly used Diebold Mariano test. The paper is organised as follows. In Section 2 we give a short overview of the relevant literature. In Section 3 the empirical data as well as the methods and tests applied will be outlined. The empirical results are presented and discussed in Section 4. Finally in section 5 we conclude the paper.

2 Literature review

There already exists a long history of both empirical and theoretical research articles dealing with the concept of forecast evaluation and the quality of professional analysts' predictions. As a matter of fact there already exist a lot of comprehensive overview articles summing up the relevant literature and methods. In this context we would like to

mention the works of Mahmoud (1984), Mincer and Zarnowitz (1969), De Gooijer and Hyndman (2006) and Hyndman and Koehler (2006). More recently Schwarzbach et al. (2014) discussed the relevance of professional financial market forecasts for investment decisions of asset managers in the life insurance industry.

The literature dealing with the assessment of the quality of survey forecasts for interest rates does also reach back for more than three decades (see for e.g., Friedman, 1980; Belongia 1987; Hafer and Hein, 1989). More recently and with a special focus on the evaluation of survey forecasts for long-term and short-term interest rates for various countries the work of Spiwoks et al. (2008, 2009, 2010) is very relevant for the focus of our paper. In their studies the authors focus inter alia on the aspects of topically orientated trend adjustments by individual forecasters. Dealing with the relevance of interest rate forecasts – both individual and survey mean forecasts – for practitioners like asset managers in German the life insurance industry the work of Schwarzbach et al. (2012) also delivers a good indication for the relevance of this topic. Additionally Baghestani (2006) focused on the accuracy of multi period forecasts for ten years. Treasury bond rates from the Survey of Professional Forecasters (SPF). The test results indicate that the forecasts are not rational but in light of the sign accuracy test results predict the direction of change correctly. Additionally Greer (2003) tested the directional accuracy of long-term interest rate forecasts issued by The Wall Street Journal's panel of economic forecasters and concluded that the forecasts in general are less accurate.

Although not in the focus of this paper we would like to mention some highly relevant articles dealing with exchange rate forecasts and aspects of financial crisis, respectively, market intervention. The accuracy of exchange rate forecasts has been discussed in different articles. Ruelke et al. (2010) examined the expectations on the yen/dollar exchange rate from the Wall Street Journal forecast poll. They found heterogeneity of the forecasts and showed the evidence of a systematic component in the forecast error. Brissimis and Chionis (2004) explored the effects of market intervention by the ECB and the Bank of Japan and found evidence for the effectiveness for these actions. Chow and Kim (2006) stated that exchange rate flexibility stabilised the interest rates in the Asian crisis only in the short-run. Devereux et al. (2012) have been questioning the association of relative consumption growth and real exchange-rate depreciations and found no such association, regardless of exchange rate regimes in 28 countries between 1990 and 2010. Hong et al. (2007) suggested to use more sophisticated time series models with higher order conditional moments, such as Markov regime-switching models, produce better forecasts than random walk based models in case of Euro-Dollar and Yen/Dollar forecasts. Neely (2009) argued that the statistical metrics cannot measure implied volatility in foreign exchange rates futures appropriately.

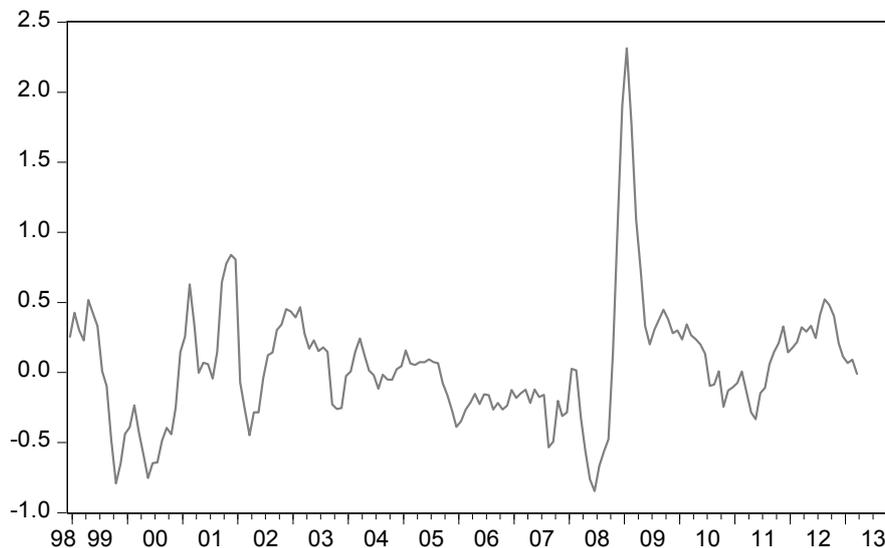
Since the outbreak of the subprime crisis the global financial crisis researchers put a lot of effort in the analysis of the crisis events themselves and the circumstances which have led to the financial turmoil (see for e.g., Basse et al., 2013; Gruppe and Lange, 2013). There are a number of papers which have been analysing the crisis effect on government bond yields (see for e.g., Gómez-Puig and Sosvilla-Rivero, 2013; Gómez-Puig and Sosvilla-Rivero, 2014). Additionally, Moro (2013) did provide an outstanding survey article dealing with that topic. Moreover, Basse et al. (2014) recently have discussed the relevance of managing the interest rate exposure of German life insurers using a stress test approach.

As the uncertainties of forecasters might rise especially during times of financial crisis also the dispersion of forecasts of individual analysts increases. To find evidence for crisis related dispersion in long-term survey forecasts for the three month EURIBOR Kunze et al. (2013) developed a simple regression model to investigate the relationship of the dispersion within the survey forecast and relevant economic variables. The work of Kunze and Gruppe (2014) is also of high relevance. The authors tested for changes of the quality of professional forecasters due to structural breaks using the Theil's U -statistic (see Theil, 1955; Theil et al., 1966).

3 Data and methodology

We want to test for the changes in the accuracy of the three months survey forecast provided by Consensus Economics Inc. for the three months EURIBOR. As a matter of fact the actual forecast horizon is four months. This is the case because the forecasters have to provide their predictions at the beginning of the months (e.g., a prediction for the three month EURIBOR at the end of March 2013 had to be provided already in December 2012). The evaluated sample ranges from December 1998 to March 2013 and because of that covers both the subprime crisis in 2007/08 and the following financial turmoil resulting in the European sovereign debt crisis (e.g., Aizenman, 2013; Basse et al., 2013). As a matter of fact the data in the first part of the series implies that there should be no structural breaks due to the common European currency area starting 1999. We focus on the relationship of the survey forecast as well as the actual three month EURIBOR rates at the end of each reporting month. The relevant forecast errors are shown in Figure 1.

Figure 1 Forecast errors (three month EURIBOR and survey forecast; in %-points)



With respect to the assessment of forecast accuracy the actual and forecasted time series have to be of the same order of integration. Otherwise the survey forecast would not fulfil one of the most elementary necessary conditions to be an adequate prediction for the three month EURIBOR. As for example Schwarzbach et al. (2012) have shown this is for example the case for the Consensus Economics Inc. survey forecasts for ten years German government bond yields with a 13 months horizon. Because of that we start our investigation by testing whether the two time series discussed above are integrated by the same order. We apply the ADF-test (see Dickey and Fuller, 1979) as well as the PP-Test (see Phillips and Perron, 1988) as rather traditional methods and the more current procedure proposed by Ng and Perron (2001). All three testing procedures indicate that the two time series considered are non-stationary and integrated by the order one (i.e. the first differences of the three month EURIBOR as well as the Survey Forecast are stationary (i.e. $I(0)$). Table 1 presents the results of the ADF-Test for the three month EURIBOR. Table 2 shows the results of the PP-test for the three month EURIBOR whereas Table 3 displays the corresponding results for the Ng-Perron-Test. Tables 4–6 show the respective test results for the Survey Forecast.

Table 1 ADF-tests three month EURIBOR and first differences

<i>Null hypothesis: three month EURIBOR has a unit root</i>		
<i>Exogenous: Constant</i>		
<i>Lag Length: 1 (Automatic – based on SIC, maxlag = 13)</i>		
	<i>t-statistic</i>	<i>Prob.*</i>
Augmented Dickey-Fuller test statistic:	-1.213900	0.6681
<i>Null hypothesis: D (three month EURIBOR) has a unit root</i>		
<i>Exogenous: Constant</i>		
<i>Lag Length: 0 (Automatic– based on SIC, maxlag = 13)</i>		
	<i>t-statistic</i>	<i>Prob.*</i>
Augmented Dickey-Fuller test statistic	-7.039561	0.0000

Table 2 PP-test three month EURIBOR and first differences

<i>Null hypothesis: three month EURIBOR has a unit root</i>		
<i>Exogenous: Constant</i>		
<i>Bandwidth: 8 (Newey-West automatic) using Bartlett kernel</i>		
	<i>Adj. t-Stat</i>	<i>Prob.*</i>
Phillips-Perron test statistic	-1.189935	0.6785
<i>Null hypothesis: D (three month EURIBOR) has a unit root</i>		
<i>Exogenous: Constant</i>		
<i>Bandwidth: 5 (Newey-West automatic) using Bartlett kernel</i>		
	<i>Adj. t-Stat</i>	<i>Prob.*</i>
Phillips-Perron test statistic	-7.132518	0.0000

Table 3 Ng-Perron-test three month EURIBOR and first differences

<i>Null hypothesis: three month EURIBOR has a unit root</i>				
<i>Exogenous: Constant</i>				
<i>Lag length: 1 (spectral GLS-detrended AR based on SIC, maxlag = 13)</i>				
<i>Sample: 1998M12 2013M03</i>				
<i>Included observations: 172</i>				
	<i>MZa</i>	<i>MZt</i>	<i>MSB</i>	<i>MPT</i>
Ng-Perron test statistics	-3.64766	-1.09823	0.30108	6.79883
Asymptotic critical values*:	1% -13.8000	-2.58000	0.17400	1.78000
	5% -8.10000	-1.98000	0.23300	3.17000
	10% -5.70000	-1.62000	0.27500	4.45000
<i>Null hypothesis: D (three month EURIBOR) has a unit root</i>				
<i>Exogenous: Constant</i>				
<i>Lag length: 1 (spectral GLS-detrended AR based on SIC, maxlag = 13)</i>				
<i>Sample: 1999M01 2013M03</i>				
<i>Included observations: 171</i>				
	<i>MZa</i>	<i>MZt</i>	<i>MSB</i>	<i>MPT</i>
Ng-Perron test statistics	-55.7109	-5.27427	0.09467	0.44855
Asymptotic critical values*:	1% -13.8000	-2.58000	0.17400	1.78000
	5% -8.10000	-1.98000	0.23300	3.17000
	10% -5.70000	-1.62000	0.27500	4.45000

Table 4 ADF-tests survey forecast and first differences

<i>Null hypothesis: survey forecast has a unit root</i>		
<i>Exogenous: Constant</i>		
<i>Lag Length: 2 (automatic – based on SIC, maxlag = 13)</i>		
	<i>t-Statistic</i>	<i>Prob.*</i>
Augmented Dickey-Fuller test statistic	-1.832663	0.3637
<i>Null Hypothesis: D (survey forecast) has a unit root</i>		
<i>Exogenous: Constant</i>		
<i>Lag Length: 1 (automatic – based on SIC, maxlag = 13)</i>		
	<i>t-Statistic</i>	<i>Prob.*</i>
Augmented Dickey-Fuller test statistic	-4.776852	0.0001

Table 5 PP-tests survey forecast and first differences

<i>Null Hypothesis: Survey Forecast has a unit root</i>		
<i>Exogenous: Constant</i>		
<i>Bandwidth: 8 (Newey-West automatic) using Bartlett kernel</i>		
	<i>Adj. t-Stat</i>	<i>Prob.*</i>
Phillips-Perron test statistic	-1.197263	0.6753
<i>Null Hypothesis: D (Survey Forecast) has a unit root</i>		
<i>Exogenous: Constant</i>		
<i>Bandwidth: 6 (Newey-West automatic) using Bartlett kernel</i>		
	<i>Adj. t-Stat</i>	<i>Prob.*</i>
Phillips-Perron test statistic	-7.162526	0.0000

Table 6 Ng-Perron-test three month EURIBOR and first differences

<i>Null hypothesis: survey forecast has a unit root</i>					
<i>Exogenous: Constant</i>					
<i>Lag length: 1 (Spectral GLS-detrended AR based on SIC, maxlag = 13)</i>					
<i>Sample: 1998M12 2013M03</i>					
<i>Included observations: 172</i>					
		<i>MZa</i>	<i>MZt</i>	<i>MSB</i>	<i>MPT</i>
Ng-Perron test statistics		-7.93342	-1.75269	0.22093	3.96142
	1%	-13.8000	-2.58000	0.17400	1.78000
Asymptotic critical values*:	5%	-8.10000	-1.98000	0.23300	3.17000
	10%	-5.70000	-1.62000	0.27500	4.45000
<i>Null Hypothesis: d(Survey Forecast) has a unit root</i>					
<i>Exogenous: Constant</i>					
<i>Lag length: 1 (Spectral GLS-detrended AR based on SIC, maxlag = 13)</i>					
<i>Sample: 1999M01 2013M03</i>					
<i>Included observations: 171</i>					
		<i>MZa</i>	<i>MZt</i>	<i>MSB</i>	<i>MPT</i>
Ng-Perron test statistics		-36.2717	-4.25151	0.11721	0.69644
	1%	-13.8000	-2.58000	0.17400	1.78000
Asymptotic critical values*:	5%	-8.10000	-1.98000	0.23300	3.17000
	10%	-5.70000	-1.62000	0.27500	4.45000

Furthermore, for the survey forecast to be a consistent prediction of the time series of the three month EURIBOR as well as the forecast time series have to be co-integrated examining I(1) variables (see also Schwarzbach et al., 2012). We tested for co-integration by examining the grade of integration of the residuals of the regression equation

displayed in Table 7. To estimate this equation we employ the fully modified least squares regression approach (see Phillips and Hansen, 1990). This is an improved version of the commonly used testing procedure proposed by Engle and Granger (1987).

Table 7 Co-integration estimation results

<i>Dependent variable: three month EURIBOR</i>				
<i>Method: fully modified least squares (FMOLS)</i>				
<i>Sample (adjusted): 1999M01 2013M03</i>				
<i>Included observations: 171 after adjustments</i>				
<i>Co-integrating equation deterministic: C</i>				
<i>Variable</i>	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-Statistic</i>	<i>Prob.</i>
Survey forecast	1.020394	0.032762	3.114.611	0.0000
C	-0.059373	0.097729	-0.607524	0.5443
R^2	0.904484	Mean dependent var		2634713
Adjusted R^2	0.903919	S.D. dependent var		1418293
S.E. of regression	0.439628	Sum squared resid		3266306
Durbin-Watson stat	0.229892	Long-run variance		0.317724

Thus the co-integrating vector seems to be $(1, -1)$. From this result it is appropriate to consider, respectively, further evaluate the forecast errors (i.e. the difference between the Survey Forecast and the Three Month EURIBOR). The resulting time series is shown in Figure 1. As it can be seen from Figure 1 there seems to be some crisis related changes in the long-run relationship between the three month EURIBOR and the corresponding forecast time series. There are some well-known problems with co-integration tests in the presence of structural breaks (e.g., Gregory and Hansen, 1996). Structural change produces a bias to reject co-integration when it is present. Thus, it is important to note that we have found co-integration among the variables examined here in spite of possible problems with structural change.

From the regression presented in Table 7 it can be assumed that the forecast error time series is stationary (i.e. $I(0)$). Because of that we apply a simple time series ARMA model. The estimation results of the ARMA model are presented in Table 8. The lag structure of the ARMA model has been determined using the Autocorrelation as well as Partial Autocorrelation functions.

As it has been proposed inter alia by Basse et al. (2009) as well as Sibbertsen et al. (2014) we search for crisis related structural breaks within this simple relationship. We apply the Quandt Andrews test to the ARMA model presented above (see Andrews, 1993; Basse et al., 2009). For this testing procedure no specific dates of possible break points have to be defined in advance – this is a clear advantage in comparison to breakpoint tests like the Chow breakpoint test which requires specific assumptions of break point dates (see Chow, 1960). As a matter of fact the Quandt Andrews test procedure is based on the Chow breakpoint test. The results of the Quandt-Andrews Breakpoint test are presented in Table 9. As Table 9 shows the most likely breakpoint is October 2008.

Table 8 Estimation results ARMA model

<i>Dependent variable: forecast error</i>				
<i>Method: least squares</i>				
<i>Sample (adjusted): 1999M02 2013M03</i>				
<i>Included observations: 170 after adjustments</i>				
<i>Convergence achieved after three iterations</i>				
<i>Variable</i>	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-Statistic</i>	<i>Prob.</i>
C	0.030931	0.076374	0.404991	0.6860
AR(1)	1.349408	0.065719	2.053.287	0.0000
AR(2)	-0.523520	0.065676	-7.971.201	0.0000
R^2	0.845293	Mean dependent var		0.040224
Adjusted R^2	0.843440	S.D. dependent var		0.438000
S.E. of regression	0.173306	Akaike info criterion		-0.650022
Sum squared resid	5.015860	Schwarz criterion		-0.594685
Log likelihood	5.825.190	Hannan-Quinn criterion		-0.627567
F -statistic	4.562.293	Durbin-Watson statistic		2.054.219
Prob (F -statistic)	0.000000			
Inverted AR roots	.67 + .26i	.67 - .26i		

Table 9 Results Quandt-Andrews breakpoint test

<i>Quandt-Andrews unknown breakpoint test</i>		
<i>Null Hypothesis: No breakpoints within 15% trimmed data</i>		
<i>Equation sample: 1999M02 2013M03</i>		
<i>Number of breaks compared: 119</i>		
<i>Statistic</i>	<i>Value</i>	<i>Prob.</i>
Maximum LR F -statistic (2008M10)	58.40647	0.0109
Maximum Wald F -statistic (2008M10)	18.81641	0.0062
Exp LR F -statistic	1.201486	0.0917
Exp Wald F -statistic	5.711429	0.0144
Ave LR F -statistic	1.941757	0.0614
Ave Wald F -statistic	5.911422	0.0577
Note: probabilities calculated using Hansen's (1997) method	-	-

In assessing the quality of the survey forecasts for the sample range December 1998 to March 2013 we apply the TOTA-coefficient proposed by Andres and Spiwoks (1999) to test for TOTA. Spiwoks et al. (2010) already have shown that bond analysts' forecasts seem to have been influenced by topically orientated trend adjustments between 1989 and 2007. The coefficient is defined as TOTA-Coefficient given in equation (1).

$$TOTA\ coefficient = \frac{R^2\text{Survey Forecast; Three Month EURIBOR}}{R^2\text{Survey Forecast; Three Month EURIBOR-h}} \quad (1)$$

where h = forecast horizon.

Following Andres and Spiwoks (1999) a value of the TOTA-coefficient below one can be seen as statistical evidence for trend adjustment behaviour. To check for such adjustments we use the R^2 from a fully modified originally least square regression (FMOLS) for both the $R^2_{\text{Survey Forecast; Three Month EURIBOR}}$ and the $R^2_{\text{Survey Forecast; Three Month EURIBOR-}h}$ because the underlying time series (i.e. the Survey Forecast and the Three Month EURIBOR as well as the Survey Forecast and the Three Month EURIBOR – h) are co-integrated. In contrast to the model specification outlined above we omit the constant from FMOLS regression for all six corresponding regressions. We go on by applying the simple but widely accepted Root Mean Squared Error (RMSE) measure (see for e.g., Hyndman and Koehler, 2006)). Additionally we evaluate the professional analysts' predictions with the help of a sign accuracy test.

And finally we use the technique of turning point analysis. This analysis might be interesting especially for asset managers because for their investment decisions correct turning point forecasts are of extremely high relevance. For a very good description of the application of a turning point analysis for interest rate forecasts see Andres and Spiwoks (2000). Finally, to check for evidence of crisis related changes in the quality of the professional forecasts we additionally evaluate the quality of the predictions over the two time horizons before and after the breakpoint in October 2008. These empirical results are presented in Section 4.

4 Empirical results and discussion

With the help of the ARMA model presented in the previous section we have been able to identify a breakpoint within the long-run relationship of the Three Month EURIBOR and the Survey Forecast in October 2008. This result is quite similar to the empirical findings of Kunze and Gruppe (2014) who found strong indication for a breakpoint in this relationship using a more basic approach. In the remainder of this section we will present the forecast evaluation results for the three resulting time horizons:

- December 1998–March 2013.
- December 1998–October 2008.
- November 2008–March 2013.

The corresponding results are presented in Table 10 (for the TOTA coefficient), Table 11 (for the RMSE) and Table 12 (for the Turning Point Analysis).

Table 10 TOTA Coefficient for the three time horizons

	<i>1999/05–2013/03</i>	<i>1999/04–2008/10</i>	<i>2008/11–2013/03</i>
TOTA coefficient	0.924	0.914	0.840

Table 11 RMSE for the three time horizons

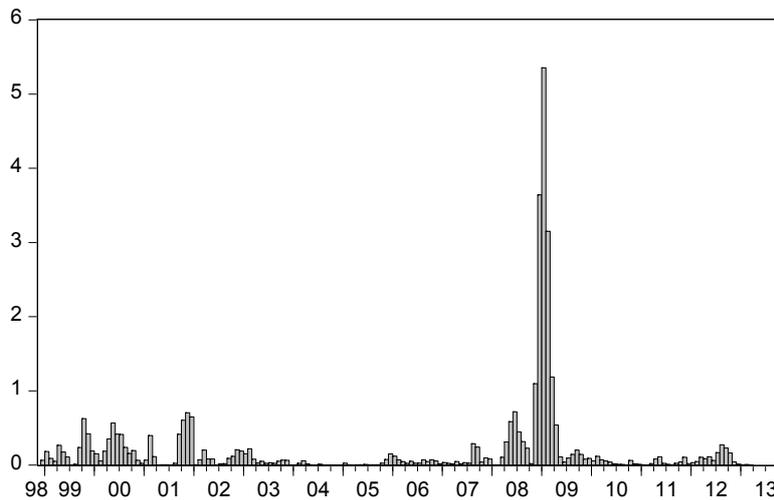
	<i>1998/12–2013/03</i>	<i>1998/12–2008/10</i>	<i>2008/11–2013/03</i>
Root mean squared error (RMSE)	0.43	0.35	0.58

Table 12 Turning Point Analysis for the three time horizons

	1999/02–2013/03	1999/02–2008/10	2008/11–2013/03
Relative error turning point forecast	0.33	0.37	0.25

The results for the TOTA-coefficient clearly point to trend adjustment behaviour. Furthermore due to the fact that especially in the time after the breakpoint the TOTA-coefficient is the lowest one might argue that trend adjustment behaviour has even become more pronounced. To further investigate these findings within future research the forecast quality matrix proposed by Andres and Spiwoks (1999) should be derived for the three relevant time horizons.

Keeping the calculation of the RMSE in mind this measure signals differences between the Three Month EURIBOR and the Survey Forecast. This means either the average deviation increased or one or more pronounced deviations between both series do exist. Analysing the calculated RMSE-results we can clearly find evidence that the quality of forecasts after the break in October/November 2008 (see section 1) tends to be lower in the average or at least in some points in time. Nonetheless too fast conclusions have to be avoided. The RMSE after the crisis might have risen after the breakpoint due to general deterioration of the forecast accuracy. But this does not have necessarily to be the case. One might even argue that the higher RMSE after the breakpoint is a result of only a few bad forecasts due to the financial turmoil in late 2008 and that after the disruption forecasters might have even been better than before the crisis. Figure 2 shows the quadratic error terms for each month (i.e. (Survey Forecast – Three Month EURIBOR)²). It is obvious that the largest error terms occur in late 2008. This observation should not come as a surprise at all since the structural break has been detected exactly in that period of time.

Figure 2 Quadratic forecast errors

We further analyse the forecast accuracy using the RMSE and apply this measure to three sequences (A, B, C):

- A: Breakpoint date \pm forecast horizon of four months (June 2008–February 2009).
- B: December 1998–May 2008.
- C: March 2009–March 2013.

The results are presented in Table 13.

Table 13 RMSE for the three time horizons

	<i>Sequence A</i>	<i>Sequence B</i>	<i>Sequence C</i>
Root mean squared error (RMSE)	1.65	0.11	0.09

The results from Table 13 point to a different conclusion than the results from Table 11: After the first months of the crisis forecasters not only have performed much better – with respect to the RMSE measure. To some extent the period after the crisis might have made it easier to predict the three month EURIBOR.

In addition to the measures presented above we also want to check for the three time horizons if the directional accuracy of the survey forecast is better than a simple random forecast (i.e. coin tossing with 50% probability). Following Kolb and Stekler (1996), we test whether the survey forecasts are independent from a random forecast using a chi square test procedure. After checking for independence we finally compare the survey forecast with the random forecast to find evidence that the directional accuracy of the survey forecast is better than a random walk forecast (see also Diebold and Lopez, 1996). The test statistics and critical values for the chi square testing procedure are presented in Table 14.

Table 14 Test statistics chi square test for the three time horizon

<i>Null Hypothesis: Survey forecast and random forecast are independent</i>			
	<i>1999/02–2013/03</i>	<i>1999/02–2008/10</i>	<i>2008/11–2013/03</i>
Test statistic	39.13	24.22	17.83
Critical Values following a chi square distribution		3.84; 5% level 6.63; 1% level	

The test statistics for all three time horizons show that the survey forecast is independent from a random forecast. Comparing the survey forecasts sign accuracy for the whole data sample with the sign accuracy of a random forecast does show that the random forecast would have outperformed the survey forecast. This can be seen from Table 15, which shows the directional forecast accuracy for the three relevant time horizons. The same is true for both the time horizon before and the time horizon after the structural break.

Table 15 Directional forecast accuracy for the three time horizon

	<i>1999/04–2013/03</i>	<i>1999/04–2008/10</i>	<i>2008/11–2013/03</i>
Correct directional forecast (survey forecast)	46	35	11
Correct directional forecast (random forecast)	86	60	26

Following Table 8 above it can be stated, that the RMSE seem to have risen as consequence of the outbreak of the financial crisis (i.e. the RMSE after the structural break in October 2008 is higher than for the measure for the time horizon before the break). Additionally, the survey forecast seems to be independent from a simple random forecast (see Table 11). With respect to the forecast accuracy we did find empirical evidence that neither before nor after the structural break the professional analysts' forecasts did outperform a random prediction. On the other hand the results from Table 12 show that the relative error of the turning point forecasts is significant smaller after the structural break (i.e. the accuracy of the forecasts with respect to the turning points in the three month EURIBOR rate is higher).

Furthermore we used a simple procedure suggested by Diebold and Mariano (1995) to test against the H_0 of equal predictive accuracy of the forecasting errors of the survey forecasts and the forecasting errors of a simple random walk model. We measure the accuracy of the two forecasting approaches using a squared loss function. The results indicate that for the whole sample the Null of the Diebold-Mariono-Test (DM-Test) cannot be rejected (p -value: 0.1428) and hence the Theil's U of 0.7455 is just a simple indication for the outperformance of the survey forecast against a naïve prediction, which is not statistically significant. For the subsample before the breakpoint the p -value is 0.06607 and the Theil's U is 0.8460556. The result points to the same interpretation as for the whole sample, at least on a confidence level of 5%. The subsample for the time horizon after the breakpoint the p value of 0.2649 does also show that the H_0 cannot be rejected on all common confidence levels. This is true although the Theil's U of 0.6793765 for itself indicate a very good performance within this sub sample. Especially the results of observing a worsening RMSE and an improving turning point forecast seem to be divergent. But this does not necessarily have to come as a surprise. The improving turning point forecast accuracy results from the fact that the ECB's central bankers had reacted to the financial turmoil and the risks of a severe economic downturn in numerous EMU member countries (see for e.g., Kunze and Gruppe, 2014) for an overview of the crisis related monetary policy of the ECB), which almost ruled out interest rate hikes. The deteriorating performance of the RMSE may in turn be a result of the unanticipated magnitude, respectively, frequency of the interest rate cuts of the ECB.

5 Conclusion

During the subprime crisis followed by financial turmoil and the European sovereign debt crisis decision makers in the financial services industry had to find quick answers for numerous problems (e.g., controlling interest rate risk or managing fixed income portfolios). In this paper we did test the accuracy of the three months survey forecast provided by Consensus Economics for the three months EURIBOR. We addressed the accuracy of the forecasts in terms of the deviation from the three months EURIBOR as well as the accuracy of the turning point forecast that is essential for strategic decisions of investors.

The most likely breakpoint for the forecasts accuracy has been found for October 2008. This result confirms that the timing of the structural change in this relationship coincides with the beginning of the financial crisis in late 2008.

Checking for quality changes before and after this breakpoint we were able to find that the forecasters' accuracy after the first months of the crisis is not much better in general. Neither before nor after the structural break the professional analysts' forecasts did outperform a random prediction. But additionally it can be stated that there is a significant improvement in the turning point forecast prediction. There is some evidence that this result is based on the fact that the ECB had to react to the financial turmoil in numerous EMU member countries, which more or less ruled out interest rate hikes.

References

- Aizenman, J. (2013) 'The Eurozone crisis: muddling through on the way to a more perfect Euro Union?', *Social Sciences*, Vol. 2, No. 4, pp.221–233.
- 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*, Vol. 219, Nos. 5/6, pp.513–542.
- Andres, P. and Spiwoks, M. (2000) *State of the art der Statistischen Ex-post-Beurteilung von Prognosen*, Sofia-Studien zur Institutionenanalyse', Vol. 00-1.
- Andrews, D.W.K. (1993) 'Tests for parameter instability and structural change with unknown change point', *Econometrica: Journal of the Econometric Society*, Vol. 61, No. 4, pp.821–856.
- Baghestani, H. (2006) 'An evaluation of the professional forecasts of U.S. long-term interest rates', *Review of Financial Economics*, Vol. 15, No. 2, pp.177–191.
- Basse, T., Friedrich, M. and Kleffner, A. (2012) 'Italian government debt and sovereign credit risk: an empirical exploration and some thoughts about consequences for European insurers', *Zeitschrift für die gesamte Versicherungswissenschaft*, Vol. 101, No. 5, pp.571–579.
- Basse, T., Friedrich, M., Kleffner, A. and von der Schulenburg, J.M.G. (2014) 'Are interest rates too low? Empirical evidence and implications for German life insurers', *Zeitschrift für die gesamte Versicherungswissenschaft*, Vol. 103, No. 1, pp.31–43.
- Basse, T., Friedrich, M. and Vazquez Bea, E. (2009) 'REITs and the financial crisis: empirical evidence from the U.S.', *International Journal of Business and Management*, Vol. 4, No. 11, pp.3–10.
- Basse, T., Reddemann, S., Riegler, J.-J. and Graf von der Schulenburg, J.-M. (2013) 'Bank dividend policy and the global financial crisis: empirical evidence from Europe', *European Journal of Political Economy*, Vol. 34, pp.S32–S39.
- Belongia, M.T. (1987) 'Predicting interest rates: a comparison of professional and market-based forecasts', *Federal Reserve Bank of St. Louis*, pp.9–15.
- Brissimis, S.N. and Chionis, D.P. (2004) 'Foreign exchange market intervention: implications of publicly announced and secret intervention for the euro exchange rate and its volatility', *Journal of Policy Modeling*, Vol. 26, No. 6, pp.661–673.
- Chow, G.C. (1960) 'Tests of equality between sets of coefficients in two linear regressions', *Econometrica*, Vol. 28, No. 3, pp.591–605.
- Chow, H.K. and Kim, Y. (2006) 'Does greater exchange rate flexibility affect interest rates in post-crisis Asia?', *Journal of Asian Economics*, Vol. 17, No. 3, pp.478–493.
- De Gooijer, J.G. and Hyndman, R.J. (2006) '25 years of time series forecasting', *International Journal of Forecasting*, Vol. 22, No. 3, pp.443–473.
- Devereux, M.B., Smith, G.W. and Yetman, J. (2012) 'Consumption and real exchange rates in professional forecasts', *Journal of International Economics*, Vol. 86, No. 1, pp.33–42.
- Dickey, D.A. and Fuller, W.A. (1979) 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association*, Vol. 74, No. 366, pp.427–431.
- Diebold, F.X. und Lopez, J.A. (1996) 'Forecast evaluation and combination', in Maddala, G.S. and Rao, C.R. (Hrsg): *Handbook of Statistics*, Vol. 14, pp.241–268.

- Diebold, F.X. and Mariano, R. (1995) 'Comparing predictive accuracy', *Journal of Business and Economic Statistics*, Vol. 13, No. 3, pp.253–265.
- Engle, R.F. and Granger, C.W.J. (1987) 'Co-Integration and error correction: representation, estimation, and testing', *Econometrica*, Vol. 55, No. 2, pp.251–276.
- Friedman, B.M. (1980) 'Survey evidence on the "rationality" of interest rate expectations', *Journal of Monetary Economics*, Vol. 6, No. 4, pp.453–465.
- Gómez-Puig, M. and Sosvilla-Rivero, S. (2013) 'Granger-causality in peripheral EMU public debt markets: a dynamic approach', *Journal of Banking & Finance*, Vol. 37, No. 11, pp.4627–4649.
- Gómez-Puig, M. and Sosvilla-Rivero, S. (2014) 'Causality and contagion in EMU sovereign debt markets', *International Review of Economics & Finance*, Vol. 33, pp.12–27.
- Gregory, A.W. and Hansen, B.E. (1996) 'Residual-based tests for cointegration in models with regime shifts', *Journal of Econometrics*, Vol. 70, No. 1, pp.99–126.
- Greer, M. (2003) 'Directional accuracy tests of long-term interest rate forecasts', *International Journal of Forecasting*, Vol. 19, No. 2, pp.291–298.
- Gruppe, M. and Lange, C. (2013) 'Spain and the European sovereign debt crisis', *European Journal of Political Economy*, Vol. 34, pp.S3–S8.
- 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*, Vol. May/June, pp.33–42.
- Hong, Y., Li, H. and Zhao, F. (2007) 'Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates', *Journal of Econometrics*, Vol. 141, No. 2, pp.736–776.
- Hyndman, R.J. and Koehler, A.B. (2006) 'Another look at measures of forecast accuracy', *International Journal of Forecasting*, Vol. 22, No. 4, pp.679–688.
- Kolb, R.A. and Stekler, H.O. (1996) 'How well do analysts forecast interest rates?', *Journal of Forecasting*, Vol. 15, No. 5, pp.385–394.
- Kunze, F. and Gruppe, M. (2014) 'Performance of survey forecasts by professional analysts: did the European debt crisis make it harder or probably even easier?', *Social Sciences*, Vol. 3, pp.128–139.
- Kunze, F., Kramer, J. and Rudschuck, N. (2013) 'Interest rate forecasts in times of financial crisis: What might be interesting to know?', *European Journal of Political Economy*, Vol. 34, pp.S45–S52.
- Mahmoud, E. (1984) 'Accuracy in forecasting: a survey', *Journal of Forecasting*, Vol. 3, No. 2, pp.139–159.
- Mincer, J.A. and Zarnowitz, V. (1969) 'The evaluation of economic forecasts. Economic forecasts and expectations: analysis of forecasting behavior and performance', *NBER*, pp.1–46.
- Moro, B. (2013) 'Lessons from the European economic and financial great crisis: a survey', *European Journal of Political Economy*, pp.1–29.
- Neely, C.J. (2009) 'Forecasting foreign exchange volatility: why is implied volatility biased and inefficient? And does it matter?', *Journal of International Financial Markets, Institutions and Money*, Vol. 19, No. 1, pp.188–205.
- Ng, S. and Perron, P. (2001) 'Lag length selection and the construction of unit root tests with good size and power', *Econometrica*, Vol. 69, No. 6, pp.1519–1554.
- Phillips, P.C.B. and Hansen, B.E. (1990) 'Statistical inference in instrumental variables regression with I (1) processes', *The Review of Economic Studies*, Vol. 57, No. 1, pp.99–125.
- Phillips, P.C.B. and Perron, P. (1988) 'Testing for a unit root in time series regression', *Biometrika*, Vol. 75, No. 2, pp.335–346.
- Ruelke, J.C., Frenkel, M.R. and Stadtmann, G. (2010) 'Expectations on the yen/dollar exchange rate – evidence from the Wall Street Journal forecast poll', *Journal of the Japanese and International Economies*, Vol. 24, No. 3, pp.355–368.

- Schwarzbach, C., Kunze, F., Rudschuck, N. and Windels, T. (2012) 'Asset management in the German insurance industry: the quality of interest rate forecasts', *Zeitschrift für die gesamte Versicherungswissenschaft*, Vol. 101, No. 5, pp.693–703.
- Schwarzbach, C., Kunze, F., Rudschuck, N. and Windels, T. (2014) 'Stock investments for German life insurers in the current low interest environment: more homework to do', *Zeitschrift für die gesamte Versicherungswissenschaft*, Vol. 103, No. 1, pp.1–19.
- Sibbertsen, P., Wegener, C. and Basse, T. (2014) 'Testing for a break in the persistence in yield spreads of EMU government bonds', *Journal of Banking and Finance*, Vol. 41, pp.109–118.
- Spiwoks, M., Bedke, N. and Hein, O. (2008) 'Forecasting the past: the case of US interest rate forecasts', *Financial Markets and Portfolio Management*, Vol. 22, No. 4, pp.357–379.
- Spiwoks, M., Bedke, N. and Hein, O. (2009) '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', *International Bulletin of Business Administration*, Vol. 4, pp.6–19.
- Spiwoks, M., Bedke, N. and Hein, O. (2010) 'Topically orientated trend adjustment and autocorrelation of the residuals - an empirical investigation of the forecasting behaviour of bond market analysts in Germany', *Journal of Money, Investment and Banking*, Vol. 14, pp.16–35.
- Theil, H. (1955) 'Who forecasts best?', *International Economic Papers*, Vol. 5, pp.194–199.
- Theil, H., Beerens, G.A.C., Tilanus, C.B. and De Leeuw, C.G. (1966) *Applied Economic Forecasting*, Revised Edition, North-Holland Publishing Company, Amsterdam, The Netherland.