Feasible estimation of the long term interest rate dynamics by nonlinear techniques

S. Fink[1] & J. Walde[2]
[1]*Treasury Department, Raiffeisenlandesbank Oberoesterreich, Austria*
[2]*Department of Statistics, University of Innsbruck, Austria*

Abstract

Due to the importance of the risk-free capital market interest rate, nearly all large economic and financial institutions deal with the analysis of its future development. Although sometimes advanced econometric methods (VAR, ECM) are used instead of or alongside the standard OLS regression approach, almost all of the work in this field deals with the basic assumption of (multi-variate, multi-equation) linear relationships between the variables. In our paper we try to find out whether nonlinearity can really be neglected. We apply artificial neural networks as a nonlinear modelling tool. Using monthly data from 1960–2005, we forecast the interest rate by means of multi-layer perceptrons (MLP). As a benchmark method we use vector autoregression models dealing with the identical dataset. The obtained results give evidence of the underlying nonlinearity of the problem. The MLP outperform the classical tools with regard to different error measures and especially in capturing the turning points of the interest curve.

Keywords: error correction model, multi-layer perceptrons, interest rate, nonlinear modelling.

1 Introduction

Due to the importance of the risk-free capital market rate, nearly all large economic and financial institutions deal with the analysis of its determinants and future development. In most cases the German 10-year Government Bond, the so-called “Bund”, is used as a benchmark for the European capital market. As the futures markets are highly liquid, the most widely used forecasting technique

www.witpress.com, ISSN 1743-3517 (on-line)
doi:10.2495/CF080051

is chart-technical analysis although e.g. Lo et al. show that it only turns out to be of use in the short term [1].

Beside these charting techniques, a wide range of econometric literature focuses on approaches that come up to the demands of scientific rigour. Part of the research focuses exclusively on long-term interest rates [2, 3], while the other part deals with analyzing the structure of the yield curve [4, 5]. The proposed independent variables in these approaches are as heterogeneous as they can be. Nevertheless, they have one thing in common: Although advanced econometric methods are used instead of or alongside standard OLS regressions, almost all the work in this field deals with the basic assumption of linear relationships between the variables. These considerations lead necessarily to the following two essential questions:

- What is the optimal set of independent variables that performs best in determining and forecasting the German government Bond rate?
- Is a linear modelling of the relationships suitable for this kind of problem or is there any additional nonlinear information?

In our work we focus on both steps of long-term interest rate modelling: the question of variable selection (and whether "common knowledge" among market participants can improve results) as well as the methodological question of linear versus nonlinear modelling. The remainder of the paper is organized as follows. In Section 2 interest theories are pictured briefly to get an idea about possible influence variables. We describe the used variable selection process by means of interviewing traders on equity and fixed income trading floors of 15 investment banks. The employed econometric models as well as the artificial neural network are presented together with the obtained results in Section 4. Finally, we summarize the main results and implications of our work.

2 Theory

In the following we try to identify the most common theoretical approaches and determinants for the risk-free Capital market rate.

In the classical capital theory the interest rate is explained by confronting demand for and supply of capital [6]. The interest rate parity [7] assumes integrated international capital (bond) markets where arbitrage will cause elimination of (international) price differences. For this reason foreign interest rate developments can be of influence on the domestic rates. The preferred habitat theory [8] shows that money market investments are a substitute for investments in the bond market. Returns may differ, but when the difference between returns gets out of equilibrium, it may induce capital movements to or from the money market, leading to changing bond prices. The preferred habitat theory acknowledges that investors and borrowers have a preference for a certain maturity (which makes this theory differ from the expectations theory of the term structure), but that changing prices in either the money market or bond market can change the investment or borrowing decisions. Hence movements in the money markets affect the long-term interest rate [9]. Fisher's interest rate theory states that investors want to be compensated for inflation [10]. They add the expected inflation over the investment period to the market clearing real interest rate. Portfolio theory asserts that the difference in asset prices is caused by differences in risk, where it is assumed that investors are risk averse and that they will only invest in higher risk assets when they are being compensated for this [11].

3 Data

To select the relevant independent variables for our problem we conducted a survey. The questionnaire contained all 18 variables proposed by theory. This questionnaire was sent to traders on equity and fixed income trading floors of 15 investment banks in the USA, Great Britain, France, Germany, and Austria. They were asked to select the variables that for them were the key factors influencing long term European government Bond yield. Furthermore they were asked about the sign of the influence and were given the opportunity of proposing additional important variables missing in the given list.

The survey resulted in the following variable set. For all variables we used end-of-month data from 1960/01/31 to 2003/12/31, the expected signs of the influences are in parentheses.

Foreign long-term interest rates (+): Traders focus on the US market as the largest foreign economic zone. As independent variable we use the average yield of the US 10 year treasury bills (USD10Y, source: Bloomberg).

Exchange rate (-) respectively (+): Our survey confirmed the importance of the exchange rate in general and the USD/EUR exchange rate in particular (ERUSDEUR, source: Federal Reserve Bank of St. Louis). In contrast to theory the sign of the exchange rate's influence was suggested to be positive.

Short-term interest rate (+): A positive relation is assumed between the long-term and the short-term interest. Corresponding to the survey's results, we use the European as well as the US short rate. For the European rate we employ the 3-month Frankfurt Interbank Offered rate (FRIBOR) from 1960 to 1998 and the EURIBOR (European Interbank Offered rate) from 1999 to 2003 (EUR3M, source: Bloomberg). The U.S. short rate is represented by the 3-month USD-Libor (USD3M, London Interbank Offered rate).

Business cycle (+): The survey resulted in the U.S. Industrial Production Index to be the most important indicator for the state of the world economy (IPUS, source: Federal Board of Governors).

Inflation rate (+) respectively (–): We use the *Consumer Price Index for All Urban Consumers: All Items* for US inflation (CPIUS, source: U.S. Department of Labor, Bureau of Labor Statistics) and *Consumer Price Index for Germany: All Items* for German Inflation (CPIGER, source: Bloomberg). We did not get a distinct sign of the influence in our survey.

Returns on assets (+): Asset returns are suggested to play an important role for the development of the long-term Bund rate. The following two stock indices were proposed by the survey: the German DAX (DAX) and the S&P 500 (SPX) which are both performance indices (source: Reuters).

As proposed before we used the German government Bond rate as dependent variable, which is the average yield of the German 10-year government bond (GDBR10Y, source: Datastream).

All data is seasonally unadjusted. We have 526 observation units as we loose the first two observations due to data preparation.

4 Specification and estimation results

As first approach we employ an error correction model (ECM). Unit root tests (Augmented Dickey-Fuller test, Phillips-Perron test) give evidence that six out of ten variables (ERUSDEUR, CPIGER, CPIUS, USD10Y, USD3M, GDBR10Y) follow a unit root process. Johansen's cointegration test indicates at a 5% significance level one cointegrating equation. Hence, we estimate the following model: $y_t = m + \alpha_1 y_{t-1} + \beta_{01} x_{t-1}^{(1)} + \beta_{11} x_{t-2}^{(1)} + \beta_{02} x_{t-1}^{(2)} + \beta_{12} x_{t-2}^{(2)} + \ldots + \gamma_1 x_{t-1}^{(6)} + \ldots + u_t$, where $x_{t-1}^{(1)}, \ldots, x_{t-1}^{(5)}$ are $I(1)$ variables ($I(1)$ denotes integrated of order 1), and $x_{t-1}^{(6)}, \ldots, x_{t-1}^{(9)}$ are $I(0)$ variables. In order to capture the entire structure of the residuals an autoregressive term and GARCH(1,1) are used. Neither the Ljung-Box-Q-statistic nor the estimated autocorrelation and partial autocorrelation function indicate any remaining structure in the residuals. AIC is -0.672 and SIC is -0.510. The detailed results for the coefficients and their significance are given in Table 2.

Table 1: Estimation results for the model. In the second column the estimates are listed and in the following column their standard errors. The value of the statistic and the corresponding *p*-value are shown in columns four and five. Additionally, the estimates for the variance equation are given.

Independent variables	**Coefficient**	**Std. Error**	***p*-value**	**Independent variables**	**Coefficient**	**Std. Error**	***p*-value**
$\hat{m}$	0.2037	0.0702	0.0037	USD10Y(-1)	0.0957	0.0335	0.0043
GDBR10Y(-1)	0.9097	0.0200	0.0000	USD10Y(-2)	-0.0917	0.0358	0.0105
ERUSDEUR(-1)	0.2941	0.1384	0.0335	USD3M(-1)	0.0047	0.0280	0.8663
ERUSDEUR(-2)	-0.2617	0.1370	0.0562	USD3M(-2)	0.01891	0.0283	0.5044
CPIGER(-1)	-0.0037	0.0205	0.8581	DAX(-1)	-0.0019	0.0012	0.1230
CPIGER(-2)	0.0256	0.0202	0.2056	IPUS(-1)	0.0058	0.0025	0.0234
CPIUS(-1)	0.0255	0.0240	0.2868	SPX(-1)	-0.0010	0.0016	0.5151
CPIUS(-2)	-0.0250	0.0236	0.2887	EUR3M(-1)	0.0216	0.0085	0.0112
$\hat{\rho}_1$	0.3928	0.0497	0.0000				
Variance Equation							
$\hat{\omega}$	0.0009	0.0004	0.0463				
ARCH: $\hat{\gamma}_1$	0.1512	0.0419	0.0003	GARCH: $\hat{\gamma}_2$	0.8359	0.0450	0.0000

For an easier interpretation we transform the estimated mean equation into the typical error correction type expression:

$$\begin{aligned}\Delta\hat{y}_t &= 0.262\cdot\Delta ERUSDEUR_{t-1} - 0.026\cdot\Delta CPIGER_{t-1} + 0.025\cdot\Delta CPIUS_{t-1} + 0.092\cdot\Delta USD10Y_{t-1}\\ &\quad - 0.019\cdot\Delta USD3M_{t-1} - 0.002\cdot DAX_{t-1} + 0.006\cdot IPUS_{t-1} - 0.001\cdot SPX_{t-1} + 0.022\cdot EUR3M_{t-1}\\ &\quad - 0.090\cdot[y_{t-1} - 2.256 - 0.360\cdot ERUSDEUR_{t-1} - 0.243\cdot CPIGER_{t-1} - 0.006\cdot CPIUS_{t-1}\\ &\quad - 0.044\cdot USD10Y_{t-1} - 0.262\cdot USD3M_{t-1}]\end{aligned}$$

The sign of the DAX is unexpectedly negative but it is not significant. ERUSDEUR, EUR3M, and USD10Y show a significant influence on the dependent variable with the expected sign. Additionally, the business cycle variable (IPUS) has a significant influence and is useful for forecasting the yield of the German 10-year government bond.

Figure 1 gives the one period ahead forecast for the out-of-sample period defined before. In addition the 95% confidence interval and the original time series, which lies except one date inside the confidence bands, are plotted.

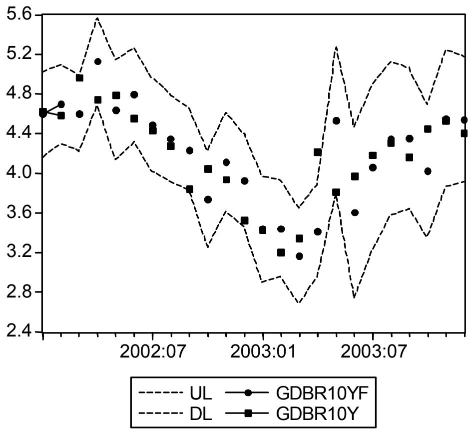

Figure 1: One-period-ahead forecasts (GDBR10YF) by means of the ECM. UL denotes the upper limit of the 95% confidence interval of the estimates and DL the down limit. The original time series (GDBR10Y) is also drawn.

The second modelling approach uses a feedforward, fully connected three layer perceptron (MLP) [12]. Although the objective of this paper is forecasting, we neglect at first the time dependence structure. This implies that the forecast of the yield of the German 10-year government bond by the MLP is solely achieved by means of the independent variables. We analyze if this procedure leads to better forecasts than the more traditional method. We split the data arbitrarily into three samples: a training set for estimating the parameters (70% of the observations), a validation set to control the optimization algorithm with respect to overfitting (20% of the observations), and a generalization set to evaluate the quality of the estimated model. The employed MLP looks like $\hat{y}_s = f\left(\boldsymbol{x}_s, \hat{\boldsymbol{\theta}}\right) = G\left(\sum_{j=1}^{HU} w_j^{(2)} \cdot G\left(\sum_{k=0}^{9} w_{kj}^{(1)} x_{k,s}\right) + w_0\right), G(x) = \frac{1}{1+e^{-x}}$, where HU is the number of hidden units, $w_{..}^{(.)}$ are the weights, $\boldsymbol{x}_s$ is the input vector of the observation units ($x_{s-1}^{(1)}, \ldots, x_{s-1}^{(9)}$), $\hat{y}_s = f\left(\boldsymbol{x}_s, \hat{\boldsymbol{\theta}}\right)$ is the network output, $\hat{\boldsymbol{\theta}}$ is the vector of all estimated parameters. For technical reasons, all input variables are transformed to the interval (-1,1) and the output variable to (0,1). The number of hidden units is varied from 1 to 15, for each constellation 50 weight initializations are carried out. The best network with respect to the mean squared error calculated on the validation set is chosen for further analysis. The so obtained optimal network achieves a determination coefficient (R^2) on the training set of 98.35%, an R^2 on the validation set of 97.16%, and an R^2 on the generalization set of 96.77%. AIC is -2.223 and SIC is -1.412, which is significantly better than the previous model.

Mainly we are interested in the forecasting ability. The previously described optimization procedure is done for all 24 networks that have to be calculated for the out-of sample period to get the one-period-ahead forecasts. Figure 2 shows the estimates obtained by the MLP (GDBR10YF) as well as the original time series (GDBR10Y). Additionally the naive forecast ($E\left(y_{t+1} \middle| \Omega_t\right) = \hat{y}_t$, where Ω_t is the set of all available information up to date t) is plotted to allow another kind of evaluation.

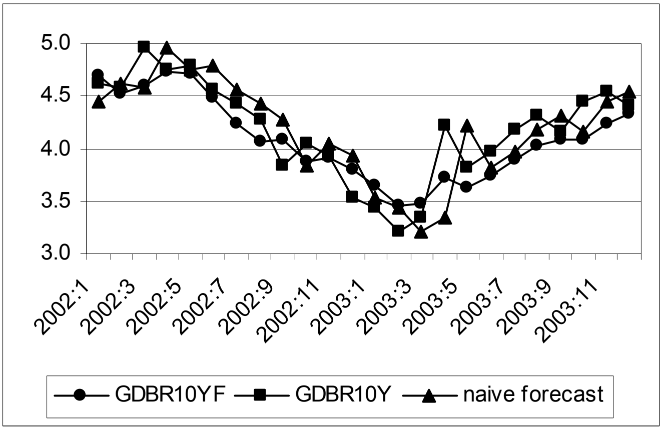

Figure 2: One-period-ahead forecasts (GDBR10YF) by means of the MLP. The original time series (GDBR10YF) is drawn as well as the estimates obtained by the naive forecast method.

The forecast by means of the MLP is quite appropriate, worth mentioning only the information contained in the independent variables is processed nonlinearly. Regarding the mean squared error (MSE, 0.052 versus 0.081), the mean absolute error (MAE, 0.194 versus 0.225), and the mean absolute percentage error (MAPE, 4.742 versus 5.488) the MLP is better than the naive forecast method. This approach is even better than the nonlinear model incorporating the remaining time structure of the residuals. The MLP accounting for the time dependency obtained a MSE of 0.069, a MAE of 0.200, and a MAPE of 4.753. To evaluate the models regarding their forecast abilities we use more forecast error statistics. The root mean squared error (RMSE) and the MAE depend on the scale of the dependent variable. Whereas the MAPE and the Theil inequality coefficient are scale invariant. The Theil inequality coefficient lies between zero and one, where zero indicates a perfect fit. To measure the ability of capturing the turning points, Theil's U_Δ statistic is used.

With respect to the forecast errors Table 2 shows the superiority of the multi layer perceptron, which do not model the time structure. Both MLP approaches do better in predicting the turning points. They are superior to the naive forecast, which obtained a Theil's U_Δ statistic of 1.137.

Table 2: Error measures of the three models.

Error Measures	ECM	MLP	MLP & time structure
RMSE	0.3191	0.2283	0.2636
MAE	0.2459	0.1938	0.2001
MAPE	6.0147	4.7393	4.7525
Theil inequality coefficient	0.0378	0.0275	0.0599
Bias proportion	0.0033	0.1929	0.1418
Variance proportion	0.0142	0.1511	0.0002
Covariance proportion	0.9825	0.6560	0.8580
Theil's U_Δ statistic	1.5768	0.7909	1.0235

5 Discussion and conclusions

As the different partial interest theories offer a wide range of influential variables on the average yield of the German 10-year government bond, we conducted a survey and asked traders on equity and fixed income trading floors of 15 investment banks. By means of the nine selected independent variables, the nonlinear modelling approaches did quite well. The performances in forecasting the two year hold-out sample of the basic structural model and of the error correction model are not distinguishable. The outstanding models are the multi-layer perceptrons that handle the relationships in a nonlinear fashion and therefore they are able to do better forecasts. Especially, the turning points of the interest curve are captured much more precisely by the MLP. This has to be due to a stable nonlinear relationship between the interest rate and the carefully selected independent variables. To be able to get reliable results by nonlinear modelling tools, sufficient data must be available. If this requirement is given and careful optimization techniques are used, this approach enhances the econometric tools for forecasting time series. The additional information can be used to enhance and enlarge the relevant information base for taking decisions. Especially turning point recognition is of practical relevance for both trading and hedging.

References

[1] Lo, A.W., Mamaysky, H. & Wang, J., Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. *The Journal of Finance*, **55 (4)**, pp. 1705–1765, 2000

[2] Breedon, F., Henry, B. & Williams, G., Long-term real interest rates: Evidence on the global capital market, *Oxford Review of Economic Policy*, **15(2)**, pp. 128–142, 1999

[3] Den Butter, F. A. C. & Jansen, P. W., An empirical analysis of the German long-term interest rate, *University of Amsterdam – Series Research Memoranda*, Research Memorandum 2000–29

[4] Diebold, F. X., Rudebusch, G. D., Boragan Aruoba, S., The Macroeconomy and the Yield Curve: A Nonstructural Analysis, CFS Working Paper Series, Number 2003/31, http://www.ifk-cfs.de/papers/03_31.pdf

[5] Ang, A., Piazzesi, M., A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics*, **50 (4)**, pp. 745–787, 2003

[6] Burmeister, E., *Capital Theory and Dynamics*, Cambridge University Press: Cambridge, 1980

[7] Blenman, L.P., Henock, L., Thatcher, J., Interest Rate Parity and the Behavior of the Bid-Ask Spread, *Journal of Financial Research*, **22 (2)**, pp. 189–206, 1999

[8] Mishkin, F. S., *The Economics of Money, Banking and Financial Markets*, 5th edition, Harper Collins: New York, 1997

[9] Riedel, F., Heterogeneous Time Preferences and Humps in the Yield Curve: The Preferred Habitat Theory Revisited, *European Journal of Finance*, **10 (1)**, pp. 3–22, 2004

[10] Phillips, P. C. B., Phillips on Fisher's Equation, *American Journal of Economics and Sociology*, **64 (1)**, pp. 125–168, 2005

[11] Nawrocki, D., Capital Market Theory: Is it relevant to practitioners? *Journal of Financial Planning*, **10 (5)**, pp. 97–103, 1997

[12] Bishop, C. M., *Neural Networks for Pattern Recognition*, Oxford University Press: New York, 1996