Money makes the world go round …

about the necessity of non-linear techniques in interest rate forecasting

Stefan Fink* & Janette F. Walde**

* Treasury Department, Raiffeisenlandesbank Oberoesterreich, Europaplatz 1a, Linz, Austria

** Department of Statistics, University of Innsbruck, Universitätsstr. 15, Innsbruck, Austria

Abstract

One of the key variables for a bank's management is the development of the "risk-free interest rate", which is the reference for all bond and loan rates as well as an indicator for the state of the economy and therefore the bank's future perspectives. Turning towards long-term analysis, the risk-free rate is usually supposed to be the return of a superior-rated government bond (in most cases the return of the German 10-year Government Bond). Due to the importance of this risk-free rate, nearly all large economic and financial institutions deal with the analysis of its future development.

In this paper we try to find out whether modelling non-linear relationships between variables can enhance forecast ability. We apply multi-layer perceptrons (MLP) as non-linear modelling tool beside an error correction model and a basic structural model with ARMA terms. Using seasonally unadjusted monthly data from 1960-2003, we forecast the interest rate for a two year hold-out sample. The obtained results give evidence of the underlying non-linearity of the problem. The MLP outperform the classical tools with regard to different error measures.

Keywords: error correction model, artificial neural networks, multi-layer perceptrons, interest rate, non-linear modeling

1 Introduction

One of the key variables for a bank's management is the development of the "risk-free interest rate". It is the reference for all bond and loan rates as well as an indicator for the state of the economy and therefore for the bank's future perspectives. Turning towards long-term analysis, which is the focus of our research, the risk-free rate is usually supposed to be the return of a superior-rated government bond.

In most cases the return of the German 10-year Government Bond, the so-called "Bund", is used for the European Capital market. Bonds with this German interest rate hold the largest market share in the total European bond market and are backed by a large and liquid futures market, of which futures are traded on both the Frankfurt and the London futures markets. For these reasons the determinants of the German rate are considered as determinants of the long-term Euro risk free rate as well.

Due to the importance of this risk-free rate, nearly all large economic and financial institutions deal with the analysis of its determinants and future development. As the futures markets are highly liquid, the most widely used forecasting technique is chart- technical analysis. These technical tools rely on the assumption that drawing "the right lines" using market price history of the Bund future can help predict future (usually short-term) market trends. In his book *A Random Walk down Wall Street*, Burton Malkiel (Malkiel et al., 1996) comes to the conclusion that "(u)nder scientific scrutiny, chart-reading must share a pedestal with alchemy." Nevertheless, technical analysis in general and "Andrew's pitchfork technique" in particular considering interest rate futures is widely used both for short- and long term trading advices although e.g. Lo et al. show that it only turns out to be of use in the short run (Lo et al., 2000).

Beside these charting techniques, a wide range of econometric literature focuses on approaches which come up the demands of scientific rigour. Part of the research focuses exclusively on (real and/or nominal) long-term interest rates (Breedon et al., 2000; Den et al., 2000), while the other part deals with analyzing and forecasting both the short and the long rates and therefore the structure of the yield curve (Diebold et al., 2003; Evans et al. 2001; Piazzesi et al., 2003). 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. Non-linearity is left out from analysis.

These considerations lead necessarily to the following two essential questions:

- What is the optimal set of independent variables which performs best in determining and forecasting the dependent variable?
- Is the modelling of the relationships linearly suitable for this kind of problem or is there any additional non-linear 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 non-linear 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 subsection 4. Finally, we summarize the main results and implications of our work.

2 Theory

Due to the importance of the risk-free long-term interest rate levels, a wide range of literature has dealt with possible determinants. Thus an even wider range of candidates for independent variables evolves from the various models. In the following we try to identify the most common theoretical approaches and determinants.

In the classical capital theory the interest rate is explained by confronting demand for and supply of capital (Burmeister, 1980). At the demand side the main actors are companies and government. The main supply stems from savings. On the bond market governments are the largest demanding party. Governments have historically built up large debts. Because of the large share that governments hold in debt markets, the change in government surplus or deficit may have a considerable impact on the long-term interest rate, at least when the share of domestic bonds in the world capital market is substantial. It can strongly affect the liquidity in the market, as has been the case in Europe and especially in the United States in 2000 (Schinasi et al., 2001). Business investments also play an important part at the demand side of the market. The investments show a strong relation with the course of business cycles (Glasner, 1997). A strong (expected) economic growth leads to higher investments, production and stock building.

The interest rate parity (Louis et al., 1999) assumes integrated international capital (bond) markets. Investors choose between investing in the home government bond or the foreign government bond. In case of price differences, arbitrage will cause their elimination. For this reason foreign interest rate developments can be of influence on the domestic rates. Furthermore, the exchange rate is of importance as it influences the net return on a foreign bond investment, $R_L = R_L^* - E^e$, where R_L is the domestic interest rate, R_L* is the foreign interest rate, and E^e is the expected appreciation of the domestic currency.

The preferred habitat theory (Mishkin, 1997) 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. For instance, if the money market interest rate rises, it increases the attractiveness to hold short - term deposits instead of long-term deposits. Investors sell bonds (bond prices decrease, effective return increases) and buy deposits in the money market. Borrowers react in an opposite way. Their preference for long-term borrowing over short-term borrowing increases, as the short-term interest rate has risen. 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. Investors are willing to change maturity when the risk premium changes. Hence movements in the money markets affect the long-term interest rate. A positive relation between the rates at the money market and the bond market is assumed, but this does not always need to be the case in practice (Riedel, 2004). For instance, a convincing response by the central bank to increase the official interest rates to fight inflation leads to a rise in short-term interest rates. Short-term inflation will be higher, while it is expected that because of the convincing central bank reaction long-term price stability is under control, which may even lead to a fall in the long term interest rate (Angeloni et al., 2001) On the other hand, an inverse yield curve is possible. In that case economic prospects for the longer run are sober, which depress new investments and reduce the demand for long capital so that long-term interest rates fall. Yet the central bank has tightened monetary policy in reaction to the overheating in the recent past so that short-term interest rates are relatively high.

The classical demand and supply theory explains the real interest rate. However, an increase in inflation will cause prices to rise on the demand and supply side, moving both the demand and supply curves to the right. Fisher's interest rate theory states that investors want to be compensated for inflation (Phillips, 1998). They add the expected inflation over the investment period to the market clearing real interest rate. Therefore $1+R_N=\left(1+R_r\right)\left(1+\pi^e\right)$, where R_N is the nominal interest rate, R_r is the real interest rate, and π^e is the expected inflation rate over maturity.

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 (Nawrocki, 1997). For instance in Sharpe's (1964) Capital Asset Pricing Model (CAPM) the assets are priced on the degree of extra risk towards the risk free rate. Acknowledging the relation between risky assets and risk free (government bonds) assets, movements in prices on, for instance, the stock market will effect bond prices and the interest rate. Evidence for this is probably the strongest in situations of financial turbulence on the stock market. Financial uncertainty on the stock market increases the risk of investing in shares. This leads to a capital flow towards a safe haven, for which usually highly liquid government bonds are used (German Bund and the US Treasury Bond; Upper, 2000). The capital flow from the stock market to the bond market reduces stock prices and causes a rise in the bond prices. Higher bond prices lower the effective interest rate.

It is not our intention to derive a formal encompassing interest theory from these five partial theories but to get theoretical foundations and guidelines to select the necessary independent variables.

3 Data

Even under consideration of the theoretical approaches, the number of candidates is almost unrestricted. For each of the theoretical driving forces, a huge number of observables could be used (GDP, unemployment, sentiment indices etc. as indicators for the state of the business cycle). 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. The participating banks are listed in the Appendix. They were asked to select the variables which 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). Japan has not been considered this important in the survey and is therefore left out.

- Exchange rate (-) respectively (+)

Theoretically, an expected appreciation of the domestic currency will attract foreign capital, which will lead to a fall in the domestic rate (negative sign). 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. The short-term interest rate is interpreted, in accordance with the preferred habitat theory, as an alternative investment choice. 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 Offer Rate (FRIBOR) from 1960 to 1998 and the EURIBOR (European Interbank Offer Rate) from 1999 to 2003 (EUR3M, source: Bloomberg). The U.S. short rate is represented by the 3 - month USD - Libor (USD3M, London Interbank Offer Rate).

- Business cycle (+)

Higher economic growth gives, according to classical capital theory, more possibilities for investment. Usually a rising business cycle goes along with rising producer and consumer confidence. 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). European or German indicators have not been considered this important for the Bund yield.

- Inflation rate (+) respectively (-)

Corresponding to Fisher's theory, both German and US inflation were considered important determinants for the risk-free rate. 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 (+)

According to the portfolio theory and confirmed by the questionnaires, 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). In order to get an impression of the time development of the dependent variable that we are going to model the yield is plotted in Figure 1.

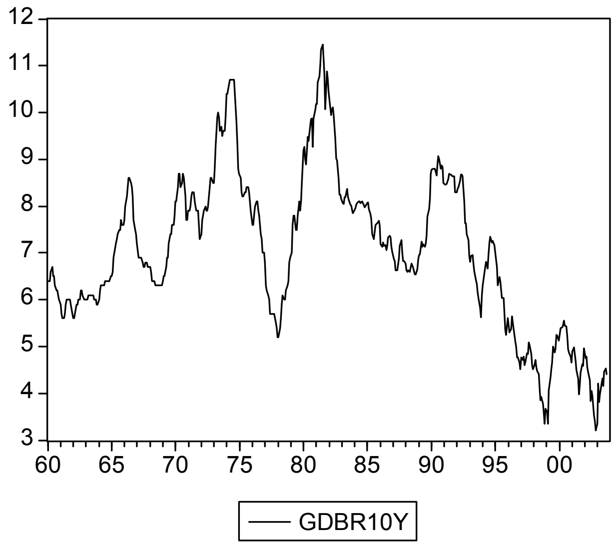

Figure 1: The monthly average yield of the German 10 - year government bond (GDBR10Y) is shown for the time period of January 1960 to December 2003.

All data is seasonally unadjusted. We have 526 observation units as we loose the first two observations due to data preparation (one date to obtain returns from the price series and the other because of one lag for forecasting).

4 Specification and Estimation Results

4.1 Basic Structural Model

In our first model (BSM) all independent variables are included linearly. To avoid dependences in the residuals, we employ three autoregressive terms. Additionally, we estimate a GARCH(1,1) model to grasp the structure of the variance. Hence, our model looks like:

$$\begin{aligned} y_t &= \alpha + \beta_1 x_{t-1}^{(1)} + \beta_2 x_{t-1}^{(2)} + \ldots + \beta_9 x_{t-1}^{(9)} + u_t \\ u_t &= \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-12} + \varepsilon_t \\ \sigma_t^2 &= \omega + \underbrace{\gamma_1 \varepsilon_{t-1}^2}_{\text{ARCH term}} + \underbrace{\gamma_2 \sigma_{t-1}^2}_{\text{GARCH term}} \end{aligned} \tag{1}$$

The estimated model is highly significant with an adjusted determination coefficient (R^2) of 98.73%. The Durbin Watson showed no further autocorrelation in the residual like the Ljung-Box-Q-statistic. The autocorrelation function as well as the partial autocorrelation function show no significant correlation at any lag. For later comparisons, the Akaike information criterion (AIC) is -0.647 and the Schwartz information criterion (SIC) -0.515. The estimated coefficients and their corresponding *p*-values are shown in Table 1.

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

Independent variables	**Coefficient**	**Std. Error**	**z-Statistic**	**p-value**
CPIGER(-1)	-0.0092	0.0214	-0.4278	0.6688
CPIUS(-1)	0.0269	0.0248	1.0856	0.2776
DAX(-1)	-0.0017	0.0009	-2.0214	0.0432
ERUSDEUR(-1)	0.3647	0.1320	2.7640	0.0057
EUR3M(-1)	0.0461	0.0148	3.1244	0.0018
IPUS(-1)	0.0010	0.0018	0.6094	0.5422
SPX(-1)	2.95E-05	0.0010	0.0290	0.9769
USD10Y(-1)	0.1081	0.0364	2.9714	0.0030
USD3M(-1)	0.0034	0.0292	0.1172	0.9067
$\hat{\alpha}$	4.8444	0.4704	10.2978	0.0000
$\hat{\rho}_1$	1.2854	0.0486	26.4552	0.0000

$\hat{\rho}_2$	-0.2853	0.050994	-5.595144	0.0000
$\hat{\rho}_3$	-0.0229	0.010648	-2.152406	0.0314
Variance Equation				
$\hat{\omega}$	0.0009	0.0004	2.1344	0.0328
ARCH: $\hat{\gamma}_1$	0.1337	0.0350	3.8206	0.0001
GARCH: $\hat{\gamma}_2$	0.8508	0.0380	22.4154	0.0000

The sign of the influence of the German DAX is unexpectedly negative, but its strength indicates that the German stock market is not that influential for international asset allocation. The low historical correlation between GDAX and other US and international indices (e.g. Dow Jones Industrial, Morgan Stanley Corporate Index) supports this presumption. The strong positive impact of the EUR/USD exchange rate underlines the survey's results and puts interest rate parity's implications into another perspective.

Both the US long term and the European short term rate confirm the supposed signs. The influence of the long term yield is twice as strong as the short rate.

The insignificance of S&P 500 data was unexpected, as portfolio-theoretical considerations consider stock returns as a main driver for the long term Bund rate.

As our main objective is forecasting, we divided the sample into the in-sample period from January 1960 to December 2001 and the out-of-sample period from January 2002 to December 2003. For the out-of-sample period we made one-period-ahead forecasts. These estimates are shown in Figure 1 together with their 95% confidence interval (UL denotes the upper limit, DL the down limit) and the original time series.

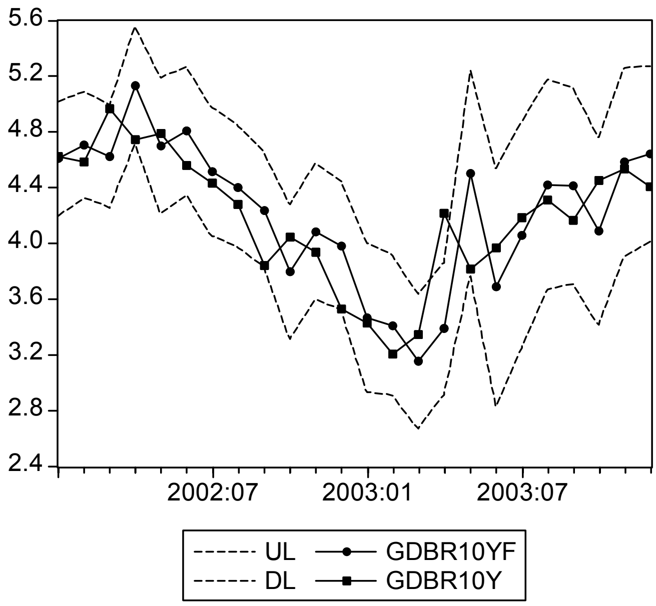

Figure 1: One-period-ahead forecasts (GDBR10YF) by means of model (1). UL denotes the upper limit of the 95% confidence interval of the estimates and DL the down limit. The original time series (GDBR10YF) is also drawn.

As Figure 1 demonstrates the forecast is quite good as all values of the original time series except one lie inside the 95% confidence interval.

4.2 Error Correction Model

As second model we employ an error correction model (ECM). Unit root tests give evidence that six (including the dependent variable) out of ten variables follow a unit root process. These are ERUSDEUR, CPIGER, CPIUS, USD10Y, USD3M, and the dependent variable. Johansen's cointegration test indicates one cointegration equation. We estimate the following model:

$$\begin{aligned} y_t = m + \alpha_1 y_{t-1} + \beta_{01} x_{t-1}^{(1)} + \beta_{11} x_{t-2}^{(1)} + \beta_{02} x_{t-1}^{(2)} + \beta_{12} x_{t-2}^{(2)} + \ldots + \beta_{05} x_{t-1}^{(5)} + \beta_{15} x_{t-2}^{(5)} \\ + \gamma_1 x_{t-1}^{(6)} + \ldots + \gamma_4 x_{t-1}^{(9)} + u_t, \end{aligned} \quad (2)$$

where $x_{t-1}^{(1)},\ldots,x_{t-1}^{(5)}$ are $I(1)$ variables, $x_{t-1}^{(6)},\ldots,x_{t-1}^{(9)}$ are $I(0)$ variables, $I(1)$ denotes integrated of order 1.

This equation can easily be written in the equivalent but more familiar form of the ECM given in equation (2').

$$\begin{aligned}\Delta y_t &= \beta_{01}\Delta x_{t-1}^{(1)} + \ldots + \beta_{05}\Delta x_{t-1}^{(5)} + \gamma_1 x_{t-1}^{(6)} + \ldots + \gamma_4 x_{t-1}^{(9)} - \\ &-\left(1-\alpha_1\right)\left[y_{t-1} - a - b_1 x_{t-1}^{(1)} - \ldots - b_5 x_{t-1}^{(5)}\right] + u_t,\end{aligned} \tag{2'}$$

where $a = \dfrac{m}{1-\alpha_1}, b_1 = \dfrac{\beta_{01}+\beta_{11}}{1-\alpha_1}, \ldots, b_5 = \dfrac{\beta_{05}+\beta_{15}}{1-\alpha_1}$.

Again to capture the entire structure of the residuals an autoregressive term and GARCH(1,1) are used. The adjusted R^2 is quite high (98.73%) and highly significant. The Durbin Watson is a little bit higher but still not significant. Neither the Ljung-Box-Q-statistic nor the estimated autocorrelation and partial autocorrelation function respectively 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 3.

Table 3: Estimation results for the model of equation (2). In the second column the estimates are listed with their standard errors in the following column. The value of the statistic and the corresponding *p*-value are shown in column four and five respectively. Additionally, the estimates for the variance equation are given.

Independent variables	**Coefficient**	**Std. Error**	**z-Statistic**	**p-value**
$\hat{m}$	0.2037	0.0702	2.9031	0.0037
GDBR10Y(-1)	0.9097	0.0200	45.3931	0.0000
ERUSDEUR(-1)	0.2941	0.1384	2.1261	0.0335
ERUSDEUR(-2)	-0.2617	0.1370	-1.9094	0.0562
CPIGER(-1)	-0.0037	0.0205	-0.1788	0.8581
CPIGER(-2)	0.0256	0.0202	1.2657	0.2056
CPIUS(-1)	0.0255	0.0240	1.0652	0.2868
CPIUS(-2)	-0.0250	0.0236	-1.0609	0.2887
USD10Y(-1)	0.0957	0.0335	2.8579	0.0043
USD10Y(-2)	-0.0917	0.0358	-2.5603	0.0105
USD3M(-1)	0.0047	0.0280	0.1684	0.8663
USD3M(-2)	0.01891	0.0283	0.6675	0.5044
DAX(-1)	-0.0019	0.0012	-1.5423	0.1230
IPUS(-1)	0.0058	0.0025	2.2674	0.0234
SPX(-1)	-0.0010	0.0016	-0.6510	0.5151
EUR3M(-1)	0.0216	0.0085	2.5350	0.0112
$\hat{\rho}_1$	0.3928	0.0497	7.9065	0.0000
Variance Equation				
$\hat{\omega}$	0.0009	0.0004	1.9928	0.0463
ARCH: $\hat{\gamma}_1$	0.1512	0.0419	3.6076	0.0003

GARCH: $\hat{\gamma}_2$	0.8359	0.0450	18.5562	0.0000

For an easier interpretation we transform the estimates into the typical error correction type expression (exclusive the estimated model for the variance):

$$\begin{aligned}\Delta\hat{y}_t &= 0.294\Delta ERUSDEUR_{t-1} - 0.004\Delta CPIGER_{t-1} + 0.026\Delta CPIUS_{t-1} + 0.096\Delta USD10Y_{t-1} \\ &+ 0.005\Delta USD3M_{t-1} - 0.002DAX_{t-1} + 0.006IPUS_{t-1} - 0.001SPX_{t-1} + 0.022EUR3M_{t-1} + \\ &+ 0.090[y_{t-1} - 2.257 - 0.360ERUSDEUR_{t-1} - 0.243CPIGER_{t-1} - 0.006CPIUS_{t-1} \\ &- 0.044USD10Y_{t-1} - 0.261USD3M_{t-1}]\end{aligned}$$

Though 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 to the previous model, the business cycle variable (IPUS) has a significant influence and is useful for forecasting the yield of the German 10 - year government bond in the expected way.

Figure 2 gives the one period ahead forecasts 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.

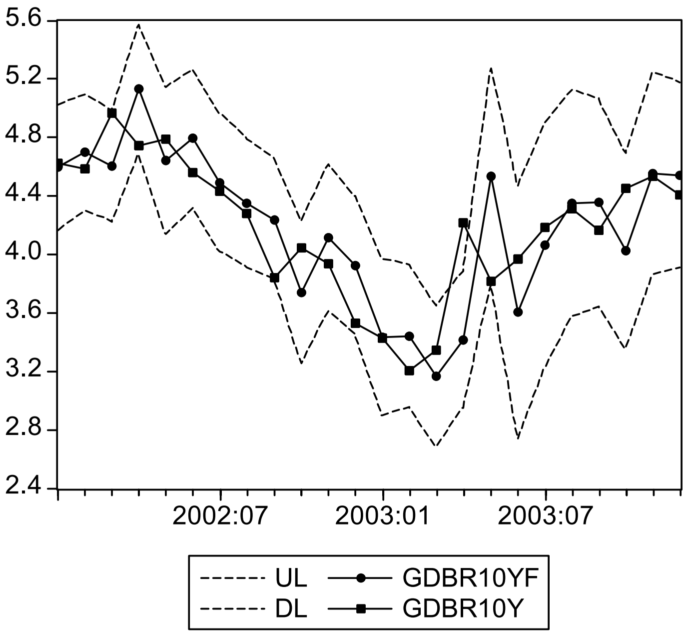

Figure 2: One-period-ahead forecasts (GDBR10YF) by means of model (2). UL denotes the upper limit of the 95% confidence interval of the estimates and DL the down limit. The original time series (GDBR10YF) is also drawn.

4.3 Artificial Neural Networks

The third modelling approach accounts for non-linear relationships between the independent variables and the dependent variable. It is somewhat more difficult to determine the non-linear influence of the independent variables, especially if a theoretically well-based assumption about the type of non-linearity is not available. For our problem a corresponding theory is not provided.

Artificial Neural Networks (ANN) are a tool for modelling non-linear functions without defining the exact functional structure. It is proved that neural networks can approximate any non-linear function with an arbitrary degree of accuracy (Hornik et al., 1998; Lek et al., 1990; Bishop, 1996). ANN are non-parametric, insensitive to the distribution of data and can generate arbitrarily complex decision boundaries (Lippman, 1987) under the requirement of sufficiently available data.

The suitability of modelling the yield of the German 10 - year government bond with the described independent variables is already shown by Fink et al. (2004). In this previous work the superiority of artificial neural networks in comparison to support vector machines is demonstrated. We rely on these results regarding the best architecture of the ANN. Hence we use a feed forward, fully connected, three layer perceptron (MLP). The calculations are carried out with the software package MATLAB Version 6.5, Release 13. Although the objective of this paper is forecasting, we employ the modelling approach of Fink et al. (2004) in the way that we neglect 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. In the modelling approach the superiority of this procedure in comparison with the econometric approaches was demonstrated (Fink et al., 2004). In this work we analyze if this procedure leads to better forecasts than the traditional methods. The data is ordered randomly. Just the dependent variable y_t stays together with the independent variables $x_{t-1}^{(1)},\ldots,x_{t-1}^{(9)}$. We split the data arbitrarily into three groups: a training data set for estimating the parameters (70% of the observations), a validation data set to control the optimization algorithm with respect to overfitting (20% of the observations), and a generalisation data set to evaluate the quality of the estimated model.

The employed MLP looks like:

$$\hat{y}_s = f\left(\mathbf{x}_s, \hat{\boldsymbol{\theta}}\right) = G\left(\sum_{j=1}^{HU} w_j^{(2)} \cdot G\left(\sum_{k=0}^{9} w_{kj}^{(1)} x_{k,s}\right) + w_0\right),$$
$$G(x) = \frac{1}{1+e^{-x}}, \tag{3}$$

where HU is the number of hidden units, $w_{..}^{(.)}$ are the weights, $\boldsymbol{x_s}$ is the input vector of observation unit s ($x_{s-1}^{(1)},\ldots,x_{s-1}^{(9)}$), $\hat{y}_s = f\left(\mathbf{x}_s, \hat{\boldsymbol{\theta}}\right)$ 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 the further analysis. The so obtained optimal network achieves a determination coefficient 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 models considering the greater number of model parameters.

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. Figure 3 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.

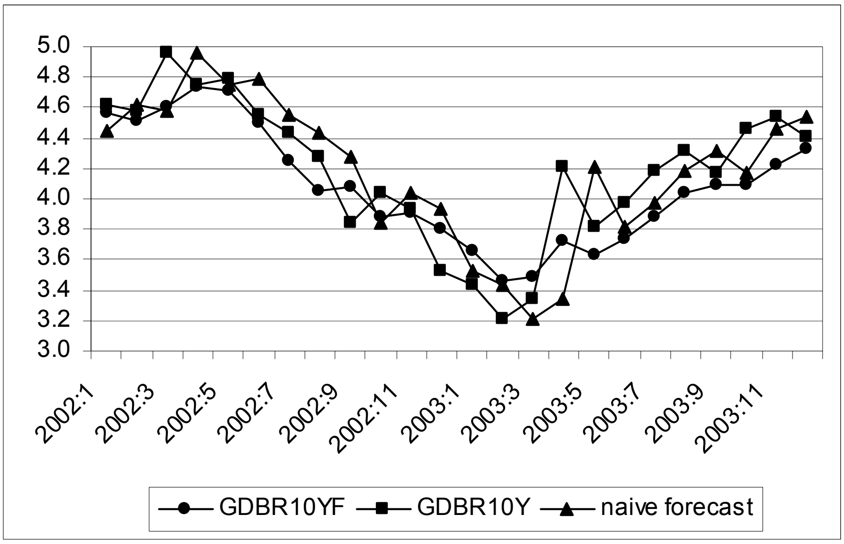

Figure 3: One-period-ahead forecasts (GDBR10YF) by means of the MLP. The original time series (GDBR10YF) is drawn as well as the estimates obtained by the naive forecast method.

The forecast by means of the MLP is quite appropriate, worth mentioning only the information contained in the independent variables is processed non-linearly. This may also be the reason that the turning points are captured well. Regarding the mean squared error (0.052 versus 0.081), the mean absolute error (0.194 versus 0.225), and the mean absolute percentage error (4.739 versus 5.488) the MLP is better than the naive forecast method. In the next subsection we give a detailed comparison of the forecast abilities of the three used approaches.

4.4 Model Comparison

To evaluate the three models regarding their forecast ability we use different forecast error statistics. The root mean squared error (RMSE) and the mean absolute error (MAE) depend on the scale of the dependent variable. The remaining two statistics, the mean absolute percentage error and the Theil inequality coefficient, are scale invariant. The Theil inequality coefficient given in equation (4) always lies between zero and one, where zero indicates a perfect fit.

$$Theil = \frac{\sqrt{\frac{1}{h}\sum_{t=T+1}^{T+h}\left(\hat{y}_t - y_t\right)^2}}{\sqrt{\frac{1}{h}\sum_{t=T+1}^{T+h}\hat{y}_t^2} + \sqrt{\frac{1}{h}\sum_{t=T+1}^{T+h}y_t^2}} \tag{4}$$

The mean squared forecast error can be decomposed into a bias proportion, variance proportion and covariance proportion:

$$
\begin{aligned}
&Bias\ proportion = \frac{\left(\frac{1}{h}\sum_{t=T+1}^{T+h}\hat{y}_t - \bar{y}\right)^2}{\frac{1}{h}\sum_{t=T+1}^{T+h}\left(\hat{y}_t - y_t\right)^2}, \\
&Variance\ proportion = \frac{\left(s_{\hat{y}} - s_y\right)^2}{\frac{1}{h}\sum_{t=T+1}^{T+h}\left(\hat{y}_t - y_t\right)^2}, \\
&Covariance\ proportion = \frac{2\left(1-r\right)s_{\hat{y}}s_y}{\frac{1}{h}\sum_{t=T+1}^{T+h}\left(\hat{y}_t - y_t\right)^2},
\end{aligned}
\tag{5}
$$

where *T* is the in-sample size, *h* the out-of-sample size, r is the correlation coefficient between y_t and $\hat{y}_t$, and *s* is the sample standard deviation.

Table 4: Error measures of the three models.

Error Measures	**BSM**	**ECM**	**MLP**
RMSE	0.3226	0.3191	0.2283
MAE	0.2548	0.2459	0.1938
MAPE	6.2263	6.0147	4.7393
Theil inequality coefficient	0.0381	0.0378	0.0275
Bias proportion	0.0184	0.0033	0.1929
Variance proportion	0.0228	0.0142	0.1511
Covariance proportion	0.9589	0.9825	0.6560

Table 4 shows the superiority of the multi layer perceptron with respect to the forecast errors.

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 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 model is the multi-layer perceptron that handles the relationships in a non-linear fashion and was therefore able to do better forecasts. Especially, the turning points between February 2002 to June 2002 and March 2003 to June 2003 are captured more precisely by the MLP. This may be due to a stable non-linear relationship between the interest rate and the independent variables (USD/EUR exchange rate, short-term interest rate, inflation rate, returns on assets, business cycle, foreign long-term interest rates).

To be able to apply MLP regarding the many degrees of freedom, 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.

This additional information can be used, alongside trader's views and charting techiques, to enhance and enlarge the relevant information base for taking decisions. Especially turning point recognition is of practical relevance for both trading and hedging, as e.g. the decision between either the bank being a fixed-interest payer or a fixed-interest receiver is the key for the bank's strategic position. Wrong positioning leads both to huge losses and an increase of Value at Risk which in turn reduces the scope of operations. Therefore, avoiding wrong decisions mean significant comparative advantage and increase the bank's prospects in all areas.

Beside the "long end" of the yield curve captured in this paper, additional supporting points of the yield curve could give relevant information about anticipated changes of the yield curve's structure and its steepness. At the moment we therefore apply MLP to a combined 2Y - 10 Y forecasting model, including 2 year (the "Bundesschatz") yields in our analysis.

References

Angeloni, I.; Kashyap, A.; Mojon, B.; Terlizzese, D. (2002): *Monetary Transmission in the Euro Area: Where do we stand?* ECB Working Paper No. 114

Bishop, C. M. (1996). *Neural Networks for Pattern Recognition*. New York: Oxford University Press

Breedon, Francis; Henry, Brian; and Williams, Geoffrey (2000): *Long-term real interest rates: Evidence on the global capital market*, Oxford Review of Economic Policy, Vol. 15 (2)

Burmeister, E. (1980): *Capital Theory and Dynamics*, Cambridge, Cambridge University Press

Upper, Christian; Worms, Andreas (2000): *Real long-term interest rates and monetary policy: a cross-country perspective*, BIS Papers No. 19, Deutsche Bundesbank

Den Butter, Frank A.C.; Jansen, Pieter W. (2000): *An empirical analysis of the German long-term interest rate,* University of Amsterdam - Series Research Memoranda, Research Memorandum 2000 - 29

Diebold, Francis X.; Rudebusch, Glenn D., Aruoba, S. Boragan (2003): *The Macroeconomy and the Yield Curve: A Nonstructural Analysis,* Federal Reserve Bank of San Francisco Working Paper 2003-18

Evans, Charles L.; Marshall, David A. (2001): *Economic Determinants of the Nominal Treasury Yield Curve*, FRB of Chicago Working Paper No. 2001-16. http://ssrn.com/abstract=294942

Fink, S. and J. F. Walde (2004): *The usefulness of econometric methods, multi-layer perceptrons and support vector machines in order to model the German interest rate*, submitted to Journal of financial and quantitative analysis

Glasner, D. (ed., 1997): *Business Cycles and Depressions*: New York: Garland Publishing

Hornik, K., Stinchcombe, M., White, H. (1989): *Multilayer Feedforward Networks Are Universal Approximators*. Neural Networks 2 (1989), 359-366

Lek, S., Delacoste, M., Baran, P., Mimopoulos, I., Lauga, J., Aulangnier, S. (1996): *Application of neural networks to modelling nonlinear relationships in ecology*. Ecol. Model. 90, 39-52

Lippman, R. D. (1987): *An Introduction to Computing with Neural Nets*, IEEE ASSP Magazine, April

Lo, Adrew.W.; Mamaysky, Harry; and Wang, Jian (2000): *Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation,* The Journal of Finance, Vol. 55 (4), August 2000

Louis, Henock; Blenman, Lloyd P.; Thatcher, Janet S. (1999): *Interest Rate Parity and the Behavoor of the Bid-Ask Spread,* Journal of Financial Research; July 15, 1999

Malkiel, Burton, 1996: *A Random Walk down Wall Street: Including a Life-Cycle Guide to Personal Investing*: W. W. Norton, New York

Mishkin, F.S. (1997): *The Economics of Money, Banking and Financial Markets*: 5th Edition, Harper-Collins, New York

Nawrocki, D. (1997): *Capital Market Theory: Is It Relevant to Practitioners?*, Journal of Financial Planning, October 1997

Phillips, C.B. (1998): *Econometric Analysis of Fisher's Equation*, No 1180 in Cowles Foundation Discussion Papers from Cowles Foundation, Yale University

Riedel, Frank (2004): Heterogeneous time preferences and interest rates--the preferred habitat theory revisited, The European Journal of Finance, Vo. 10 (1)

Schinasi, G.J.; Kramer, C.F.; Smitz, T.R. (2001): *Financial Implications of the Shrinking Supply of U.S. Treasury Securities*, International Monetary Fund

Upper, C. (2000): *How Safe Was the "Safe Haven"? Financial Market Liquidity during the 1998 Turbulences*, Discussion paper 1/00, Deutsche Bundesbank)

Appendix

Participating Banks in Survey (in alphabetical order)
Bank of America, London
BHF ING Bank, Frankfurt
BNP Paribas, Paris
CDC IXIS, Paris
CIBC, London
Citibank, London
Commerzbank, Frankfurt
Deutsche Bank, Frankfurt
Hypo Vereinsbank, Munich
JPMorgan, N.Y.
Lehman Brothers International, London
Morgan Stanley, N.Y.
Raiffeisen Landesbank Oberoesterreich, Linz
Raiffeisen Zentralbank, Wien
Societé Génerale, Paris