Probabilistic Forecasting of U.S. Treasury Bills

Diploma thesis presented to the

Institute of Applied Mathematics
Chair for Mathematical Statistics
Ruprecht Karls University of Heidelberg

by

Sam Doerken

August 11, 2011

Abstract

A set of forecasts of historical 3-month Treasury bill rates are composed and evaluated using standard performance measures and so-called profit rules. An investigation of the correlation between the two types of scoring functions reveal a stronger relationship than was suggested by previous studies. In addition, a type of probabilistic forecasts are introduced which are derived from given point forecasts. While the general benefits of such density forecasts are explained, results in this paper reveal only partial improvement of the probabilistic forecasts over their counterpart point forecasts.

Key words and phrases: U.S. Treasury bills, futures, scoring functions, probabilistic forecasts

Contents

1. List of symbols and abbreviations

$\mathbf{1}\{\dots\}$	Indicator function with condition $\{\dots\}$
n_k	Number of predictions by a forecast with horizon k
x_t	Spot rate of a T-Bill at time t
$\hat{x}^t_{t+k}$	Prediction issued at time t for the spot rate of a T-Bill at time $t+k$
$\hat{\boldsymbol{x}}^t$	Set of predictions issued at time t: $\hat{\boldsymbol{x}}^t = (\hat{x}^t_{t+1}, \dots, \hat{x}^t_{t+9})$
y^t_{t+k}	Futures rate for a T-Bill at time t with settlement at $t+k$
$\hat{\boldsymbol{y}}^t$	Set of futures rates at time t: $\hat{\boldsymbol{y}}^t = (\hat{y}^t_{t+1}, \dots, \hat{y}^t_{t+9})$
AR$_n$(r)	Autoregressive forecast with a lag of r and a rolling training period n
DA	Directional accuracy
Ecp	Error-corrected probabilistic [forecast]
MAE	Mean absolute error
Profit A	Profit rule A
RMSE	Root-mean-squared error
T-Bill	3-month U.S. Treasury bill

2. Introduction

The forecasting of government bonds, let alone tradable securities, is a substantially studied subject with promises of profitable ventures where predictions can be successfully made. While the study of forecasts is not new, relatively little attention has been paid to the considerations of scoring functions in use to express the merits of a forecast. Diebold and Li (2006) give an extensive overview of existing methods for predictions of the term structure of interest rates. On one hand, there exist traditional scoring functions that are widely used such as the root-mean-squared error which provide an admissible accuracy measure of a forecast. On the other hand, in the context of economic variables such as the price of U.S. Treasury bills, a conceivable performance measure is the gathering of potential profit that could be garnered from a forecast. The best performance measure to use depends on the aim and the context of any forecast. However, regardless of the choice, this paper aims to compare and evaluate both traditional scoring functions and ones that measure hypothetical profits for a set of forecasts of historical U.S. Treasury bill rates.

The focus of this paper is not to devise methods that produce the most accurate of forecasts. Rather, a set of forecasts is taken as given, making use of relatively naive forecasting methods that are easily reproduced and put into practice. Nevertheless, the comparison of performances thereof should provide insights into the nature of scoring functions.

Additionally, this paper seeks to introduce probabilistic forecasts as a feasible alternative to mere point forecasts. The limitations are none because a point forecast can still be derived from any probabilistic forecast although care has to be practiced to use a correct functional. Additionally, probabilistic forecasts provide information about the spread of the predicted quantity. This paradigm seems to be unevenly distributed among the sciences, with meteorology relying heavily on probabilistic forecasts while they are relatively rare in economic forecasting. In part, this may be explained by the need of point forecasts to aid in the decision-making process that is underlying many economic problems. In the end, forecasts are supposed to aid in making decisions and should therefore be evaluated in the decision making context in which they are used (Pesaran & Skouras, 2002; Granger & Pesaran, 2000). Instances where probabilistic forecasting is found in the economic sciences is e.g. Abramson and Finizza (1995) who utilize probabilistic forecasts for predictions in the market for crude oil. Similarly, Önkal and Muradoglu (1994) make use of probabilistic forecasting for predictions in the stock market. Lastly, Britton *et al.* (1998) describe the Bank of England's approach in using probabilistic forecasts for its inflation rate predictions.

While the study by Diebold and Li (2006) is comprehensive in examining forecasting techniques of interest rates, evaluation thereof is without the context of potential profits. Such an attempt was made by Leitch and Tanner (1991). In their study, Leitch and Tanner gathered historical rates of 3-month U.S. Treasury bills from 1982 up to 1988 and subjected them to various forecasting techniques. At the same time, they compared forecasts to prevailing futures rates of 3-month U.S. Treasury bills. According to every prediction, a futures contract can hypothetically be bought or sold, resulting in profits or losses depending on realizations of rates. Having established such profit rules, Leitch and Tanner then go on to compare those results with those of traditional performance measures.

This paper widely follows the approach by Leitch and Tanner: monthly forecasts of historical 3-month U.S. Treasury bill rates are conducted for the period from 1982 until end of 1996 which includes the period studied by Leitch and Tanner. Whenever possible, similar forecasting techniques are employed. Forecasts are evaluated using the same scoring functions and the correlations between performance measures are studied as well.

Chapter 3 introduces U.S. Treasury bills and futures contracts to the unfamiliar reader. It provides an overview of the studied data and establishes notation used throughout the paper. Chapter 4 explains the forecasting methods employed and introduces probabilistic forecasts in the context of given point forecasts. Chapter 5 displays the overview of the main results. In the appendices A to D, detailed description of results are provided as well as the R code used in obtaining the results.

3. U.S. Treasury bills

3.1. About Treasury bills

A 3-month U.S. Treasury bill (T-Bill) is a type of financial instrument which represents a legal promise by the U.S. Treasury Department to make a payment at a specified date in the future. The amount of the payment is called the *face value*. The time until the payment is to be made is called the *maturity* of a T-Bill, hence 3-month U.S. Treasury bills have a maturity of 3 months. They provide a risk-free investment opportunity for buyers while providing the U.S. government a means for borrowing money. They are sold at weekly held auctions, referred to as the primary market, and are up for free trade thereafter on exchanges, known as the secondary market.

Treasury bills are thought of as being the least risky form of investment available given that the full faith and credit of the U.S. government backs these securities. Since the U.S. government relies on its ability to borrow money, paying its obligations has highest priority to maintain its top credit rating. Furthermore, default can theoretically always be avoided by the administration's ability to print money, albeit at the cost of devaluating the currency.

Investor can either hold the Treasury bill until maturity, at which time the face value becomes due; or the T-Bill may be sold in the secondary markets prior to maturity. In the latter case, the investor recovers the market value of the T-Bill.

Treasury bills emerged in the wake of World War I, when the U.S. government faced difficulties in borrowing money from other countries to finance the war. The intention was to transfer debt to citizens willing to lend money and repay them during time of economic recovery (Garbade, 2008).

The daily rate of each month's last trading day was gathered, beginning on December 31, 1981 and ending on December 31, 1996. The rates are freely available at the website of the U.S. Federal Reserve Statistical Release which includes quotations of 3-month U.S. Treasury bill rates on the auction high market and secondary market (FRB, 2010). The auction high market corresponds to the emission of Treasury securities directly from the Treasury Department to buyers in auctions held every Monday. The secondary market corresponds to all open trade among investors after the security has been emitted. Since the latter trades every weekday, rates from the secondary market were chosen for this study as they seemed more feasible for a simulation of trading.

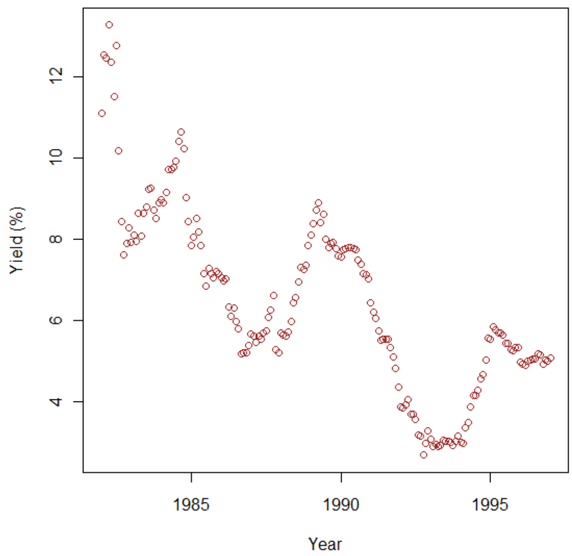


Figure 3.1.: End-of-month rates of 13-week T-Bills from 1982 to 1996

3.2. Pricing Treasury bills

The value of a Treasury bill is usually quoted by its *yield* or *rate*. The rate of an investment is the profit of an investment, meaning the sum of all future cash flows minus its price, expressed as a percentage of the face value. For example, a $100 investment resulting in a repayment of $102 has a rate of 2%. When scaled to 360 days it becomes an *annualized rate* (Mankiw, 1997): [1]

$$\text{Rate}(\%) = \frac{\text{Face Value - Price}}{\text{Face Value}} * \frac{360}{\text{Time to Maturity (days)}} * 100.$$

The terms "price" and "rate" of a T-Bill can therefore be used interchangeably as one can be derived from the other.

As with any freely traded good, the price of a T-Bill is determined by the equilibrium of supply and demand forces (Hull, 2006). In the case of a financial instrument, the equilibrium price is determined by the buy and sell orders on the exchange. All other things being equal, prices rise as supply decreases or demand increases, and fall as supply increases or demand decreases.

Factors affecting the supply of T-bills include (Cecchetti, 2006)

[1] The purpose of an *annualized rate* is to allow for comparison of rates between investments of different maturities. This is done by multiplying the rate by a factor $\frac{360}{\text{Time to Maturity (days)}}$. Though not being precise because it neglects compound interest, this is common practice as the difference is negligible.

- **Changes in government borrowing**: The government's need for borrowing money determines how many new T-Bills will be issued, changing the supply of T-Bills on the market.
- **Expected inflation**: Issuers of T-Bills care about the *real* cost of borrowing, that is, taking inflation into account. The cost of borrowing is the difference between the present value of the T-Bill and its price. Inflation lowers the present value of a future payment, thus for the issuer of a T-Bill, an inflation lowers the cost of borrowing, hence encouraging borrowing.

Demand factors include

- **Wealth**: An increase in wealth increases demand for investment, in turn increasing demand for T-Bills.
- **Alternative investments**: Changes in investment alternatives such as stocks will affect attractiveness of T-Bills to investors.
- **Taxes**: Interest earned from T-Bills is subject to federal income tax, thus change in legislation affects demand.
- **Expected inflation**: As inflation reduces the cost of borrowing, it follows that it increases the cost of lending as real value of earned interest is diminished.
- **Expected future interest rates**:[2] Investors care about future interest rates because through the concept of *arbitrage*, interest rates determine the current worth of a future payment. Assuming there is no risk involved, the price P of a security that entitles to a payment of \$X at maturity could alternatively be invested at the risk-free interest rate i. [3] Through the alternative one would receive $P(1+i)$ at maturity. Since neither option should have any advantage over the other, it is a necessary condition for price P to be such that the resulting revenue from both options be the same. Hence an inverse relationship between prices and interest rates exists:

$$\begin{aligned} &P(1+i) > \$X \Rightarrow \text{investor has no incentive to buy the security,} \\ &P(1+i) < \$X \Rightarrow \text{borrower has no incentive to sell the security,} \\ \Rightarrow\ &P(1+i) = \$X. \end{aligned}$$

The last factor will be the focus of this study as it has both a large influence on price movements of T-Bills and is very fluctuating.

[2] Note that what matters is *expected* inflation and *expected* future interest rates. Thus if one wishes to predict movements, one need not predict nominal changes of inflation or interest rates but rather the *change of people's expectations* about them.

[3] This interest rate is the theoretical interest rate at which money can be borrowed or loaned without risk.

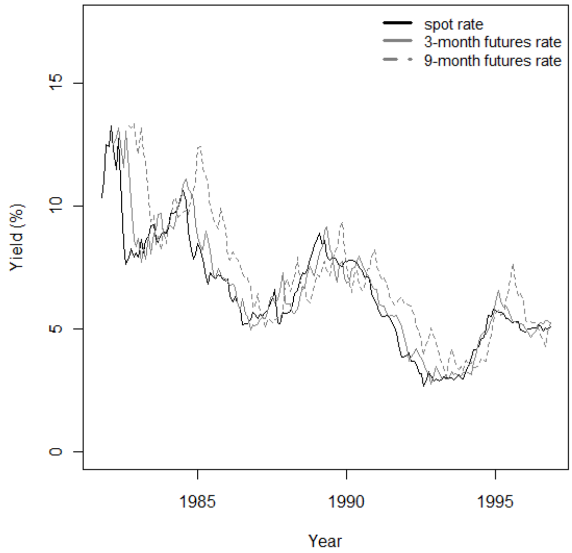


Figure 3.2.: End-of-month futures rates of T-Bill contracts with a settlement in 3 and 9 months, 1982 to 1996

Realized T-Bill rates are denoted by x_t where the lower index denotes the time from which the rate was taken. The data set thus takes the form $x_1, x_2, \ldots, x_{180}$ where x_1 is the rate observed on December 31, 1981, and x_{180} the rate from the last trading day of December 31, 1996. The data are visualized in Figure 3.1.

3.3. Futures contracts

A *futures contract* is an agreement between two parties to exchange an asset at a specific future date at a predetermined price. Futures contracts are standardized, allowing them to be traded at futures exchanges (Labuszewski & Sturm, n.d.). They are a type of financial derivative because their value is determined by price movements of the underlying asset. Using common terminology, the party agreeing to buy the asset at the settlement date is said to assume a ”long position” while the party agreeing to sell the asset is assuming a ”short position” (Hull, 2006). Futures rates will be of interest in this study because they allow for profit making with the speculation of price movements of T-Bills.

The data for the tradable futures contracts on T-Bills was purchased from the Chicago Mercantile Exchange (CME, 2010). It includes complete daily end-of-day closing prices of T-Bill contracts with a settlement within 9 months from 1982 up to year-end 1996. The last closing prices of the last trading day of every month was used. Futures rates with settlements in 3 months and 9 months are visualized in Figure 3.2 along with the spot rate, that is, the current rate, of the T-Bills.

A futures contract is denoted by y^t_{t+k} where the upper index indicates the month during which the futures contract is traded and the lower index indicates the month during which the futures contract is settled. Future contracts are settled only in the months of March, June, September, and December. If t represents the last trading day of the month and $t = 0$ corresponds to December 1981, $t = 1$ to January 1982 and so on, month-end's futures rates can be expressed in vector form as

$$\begin{aligned}
\boldsymbol{y}^1 &= (y^1_3, y^1_6, y^1_9)' \\
\boldsymbol{y}^2 &= (y^2_3, y^2_6, y^2_9)' \\
\boldsymbol{y}^3 &= (y^3_6, y^3_9, y^3_{12})' \\
\boldsymbol{y}^4 &= (y^4_6, y^4_9, y^4_{12})' \\
&\vdots \\
\boldsymbol{y}^{177} &= (y^{177}_{180}) \\
\boldsymbol{y}^{178} &= (y^{178}_{180}) \\
\boldsymbol{y}^{179} &= (y^{179}_{180}).
\end{aligned} \tag{3.1}$$

In order to evaluate forecasts, hypothetical futures rates that lie in between settlement dates are necessary. Leitch and Tanner (1991, p. 584) did this as follows:

> ”The futures rate forecasts were derived from the historical prices of the four nearest Treasury-bill future contracts. For forecasts nearer than the nearest contract, we interpolated between the current spot rate and the rate implied by the nearest contract. For forecast dates between contract dates, we interpolated between the implied forecasts of the adjoining contract dates.”

Following this method, interpolations are made as follows: if e.g. $t+k$ is such that $t+k$ is 2 month from the next settlement of a futures contract, then

$$y^t_{t+k} = y^t_{t+k-1} + \frac{1}{3}(y^t_{t+k+2} - y^t_{t+k-1}).$$

Similarly, if $t + k$ is such that it is 1 month away from the settlement, then

$$y^t_{t+k} = y^t_{t+k-2} + \frac{2}{3}(y^t_{t+k+1} - y^t_{t+k-2}).$$

With this assumption, the data set of futures prices is expanded to

$$
\begin{aligned}
\boldsymbol{y}^1 &= (y_2^1, y_3^1, \ldots, y_{10}^1)' \\
\boldsymbol{y}^2 &= (y_3^2, y_4^2, \ldots, y_{11}^2)' \\
\boldsymbol{y}^3 &= (y_4^3, y_5^3, \ldots, y_{12}^3)' \\
\boldsymbol{y}^4 &= (y_5^4, y_6^4, \ldots, y_{13}^4)' \\
&\vdots \\
\boldsymbol{y}^{177} &= (y_{178}^{177}, y_{179}^{177}, y_{180}^{177})' \\
\boldsymbol{y}^{178} &= (y_{179}^{178}, y_{180}^{178})' \\
\boldsymbol{y}^{179} &= (y_{180}^{179}),
\end{aligned}
$$

so that every forecast can be matched with a futures contract y^t to allow for a decision every month whether to buy or sell such a contract.

4. Forecasts

4.1. Point forecasts

For every last trading day of the month, starting on December 31, 1981, forecasts of the T-Bill rates are made for the last trading day for each of the subsequent 9 months. The last forecast is made on November 27, 1996 which only requires one prediction for December 31, 1996. Thus every forecast makes a total of 1584 predictions: 108 for every year except the last where only 72 predictions are required. However, three predictions had to be discarded because there was no available futures rate to match, as is necessary for calculating the later mentioned profit rules. The predictions in question without a matching futures rate were made on December 31, 1981 for January 29, April 30 and July 30, 1982, lowering the total number of predictions to 1581.

Single predictions of the T-Bill rates are denoted by $\hat{x}^t_{t+k}$ where the upper index denotes the time when the forecast is made and the lower index the time for which the rate is to be predicted.

If t represents the last trading day of the month and $t = 0$ corresponds to December 1981, $t = 1$ to January 1982 and so on, a forecast's monthly predictions can be expressed in vector form as $\hat{\boldsymbol{x}}^t = (\hat{x}^t_{t+1}, \hat{x}^t_{t+2}, \ldots, \hat{x}^t_{t+9})'$. The entire set of predictions thus takes the form of

$$
\begin{aligned}
\hat{\boldsymbol{x}}^0 &= (\hat{x}^0_1, \hat{x}^0_2, \ldots, \hat{x}^0_9)' \\
\hat{\boldsymbol{x}}^1 &= (\hat{x}^1_2, \hat{x}^1_3, \ldots, \hat{x}^1_{10})' \\
\hat{\boldsymbol{x}}^2 &= (\hat{x}^2_3, \hat{x}^2_4, \ldots, \hat{x}^2_{11})' \\
&\vdots \\
\hat{\boldsymbol{x}}^{177} &= (\hat{x}^{177}_{178}, \hat{x}^{177}_{179}, \hat{x}^{177}_{180})' \\
\hat{\boldsymbol{x}}^{178} &= (\hat{x}^{178}_{179}, \hat{x}^{178}_{180})' \\
\hat{\boldsymbol{x}}^{179} &= (\hat{x}^{179}_{180}).
\end{aligned}
$$

The following subsections introduce the four forecasting methods employed in this study.

4.1.1. Naive no-change forecast

The *naive no-change forecast* is made by setting

$$\hat{\boldsymbol{x}}^t = (x_t, x_t, \dots, x_t)',$$

i.e. by predicting that the T-Bill rate for the following nine months will be the same as the current month's.

4.1.2. Constant rate of change forecast

The *constant rate of change forecast* is done by setting

$$\hat{\boldsymbol{x}}^t = (x_t + (x_t - x_{t-1}), x_t + (x_t - x_{t-2}), \dots, x_t + (x_t - x_{t-9}))',$$

i.e. by predicting that the difference to the current spot rate since k months ago will be the same difference from the spot rate to the forecast horizon k. A conceivable alternative version where changes from the month previous to the current month are observed and the difference is linearly extrapolated for each forecast horizon was also considered. However, this method produces largely exaggerated forecasts for long horizon predictions and thus was not employed.

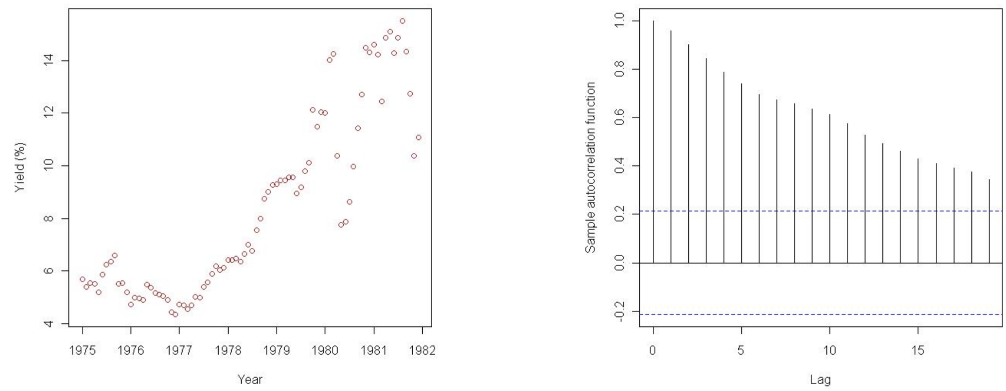
4.1.3. Autoregressive forecast

The *autoregressive forecast* is done by continuously fitting either an $AR(1)$ or $ARIMA(1,1,0)$ model to a rolling training period of the data and then make predictions based on the models. This method follows the general approach put forth by Brockwell and Davis (2006, p. 273):

> "If the data (a) exhibits no apparent deviations from stationarity and (b) has a rapidly decreasing autocorrelation function, we shall seek a suitable ARMA process to represent the mean-corrected data. If not, then we shall first look for a transformation of the data which generates a new series with the properties (a) and (b). This can frequently be achieved by differencing, [...]"

Introducing notation, an $ARMA(p,q)$ process is a stochastic process $\{X_t | t = 0, \pm 1, \pm 2, \dots\}$ where $\{X_t\}$ is stationary and for every t,

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q},$$



(a) End-of-month rates of T-Bills (b) Sample autocorrelation function

Figure 4.1.: Undifferenced data of the training period, 1975 through 1981

where $\{Z_t\} \overset{\text{iid}}{\sim} N(0, \sigma^2)$, $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q \in \mathbb{R}$. $\{X_t | t = 0, \pm 1, \pm 2, \ldots\}$ is an ARMA(p,q) process with mean μ if $\{X_t - \mu | t = 0, \pm 1, \pm 2, \ldots\}$ is an ARMA(p,q) process.

It is common to define a *backward shift operator* B such that $BX_t = X_{t-1}$. If $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 z + \ldots \theta_q z^q$, then an ARMA(p,q) process can be rewritten as

$$\phi(B)X_t = \theta(B)Z_t.$$

Depending on the stationarity of the period for each modeling, either an $AR(1)$ or $ARIMA(1,1,0)$ model is fitted. If d is a non-negative integer, then X_t is said to be an $ARIMA(p,d,q)$ process if

$$Y_t := (1 - B)^d X_t$$

is an $ARMA(p,q)$ process. The added factor $(1 - B)$ is a form of differencing to eliminate a trend component from the data. Differencing is effective for eliminating a trend component m_t if it is assumed that the data fits a process of the form

$$X_t = m_t + Y_t,$$

where Y_t is a stationary process and $m_t = \sum_{j=0}^{k} a_j t^j$ a polynomial of some unknown degree k. If a difference operator ∇ is defined as $\nabla \equiv 1 - B$ it follows that if m_t is a polynomial of degree k, then ∇m_t is of degree $k - 1$. Thus by applying the difference operator k times to the model $X_t = m_t + Y_t$, a stationary process with mean $k!a_k$ is obtained:

$$\nabla^k X_t = k!a_k + \nabla^k Y_t.$$

The initial training period from 1975 to 1982 appear to have a trend component as can be inferred from Figure 4.1. Applying the difference operator once generates a process much more indicative of stationarity as seen in Figure 4.2. Note the changing volatility of the process which is characteristic for a heteroskedastic processes.

Table 4.1.: Critical values of the KPSS test

Level	0.10	0.05	0.025	0.01
Critical value	0.119	0.146	0.176	0.216

To automate the decision process of whether or not the data from each training period should be differenced before modeling, the *Kwiatkowski-Phillips-Schmidt-Shin test* (KPSS) is used. The KPSS tests for the null hypothesis that a time series is stationary (Kwiatkowski *et al.*, 1992). The assumption is that the series can be decomposed into a deterministic trend, a random walk, and a stationary error:

$$y_t = \xi t + r_t + \epsilon_t.$$

Here, r_t is the random walk with $r_t = r_{t-1} + u_t$ where $u_t \overset{\text{iid}}{\sim} N(0, \sigma^2)$, and $\epsilon_t \overset{\text{iid}}{\sim} N(0, \sigma_\epsilon^2)$. Under the null hypothesis $\sigma^2 = 0$, y_t is trend stationary.

The test statistic is a one-sided statistic of the form

$$\frac{\sum_{t=1}^{T} S_t^2}{\hat{\sigma}_\epsilon^2} = \frac{\sum_{t=1}^{T} \left(\sum_{i=1}^{t} e_i\right)^2}{\frac{1}{T}\sum_{i=1}^{T} e_i^2},$$

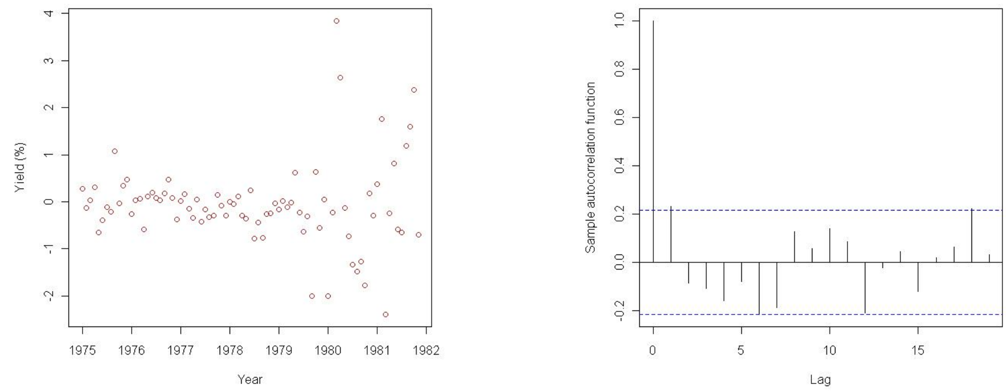
where $e_t = x_t - (r_0 + \xi t)$ are the residuals of regressing the data on an intercept and time trend, and σ_ϵ^2 is the estimate of the error variance from this regression.

The critical values of the distribution of the statistic are displayed in Table 4.1. It is taken directly as presented by Kwiatkowski *et al.* (1992).

The test for stationarity for each training period is done at the 5% confidence level.

Being able to judge whether to fit an $AR(1)$ or $ARIMA(1,1,0)$ model it follows that the parameters must be estimated. A prominent selection criteria is *Akaike's information criterion corrected* ($AICc$). It chooses parameter values that maximize a likelihood function while also assigning a cost to the introduction of each additional parameter. Although the model estimation performed here is simplified because only two parameters ϕ_1 and σ^2 are needed, the AICc is generally advocated by Brockwell and Davis (2006) as the "prime criterion for model selection". As a result, ϕ_1 and σ^2 are to be chosen such that the likelihood function

$$L\left(\phi_1, \sigma^2\right) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \left(\prod_{i=0}^{n-1} r_i\right)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\sigma^{-2}\sum_{j=1}^{n}\frac{(x_j - \phi_1 x_{j-1})^2}{r_{j-1}}\right]$$



(a) Differenced, end-of-month rates (b) Sample autocorrelation function

Figure 4.2.: Once differenced data of the training period, 1975 through 1981

is minimized, where $r_i = \frac{(x_{i+1}-\phi_1 x_i)^2}{\sigma^2}$.[1]

Having estimated ϕ_1 and σ^2 the next step is to find a best linear predictor for x_{t+k}. "Best" refers to a predictor that minimizes $\mathbb{E}[|x_{t+k} - \hat{x}^t_{t+k}|^2]$. Such a predictor is determined by a projection of x_{t+k} onto $\hat{x}^t_{t+k} \in span(x_1, \ldots, x_t)$, which requires that

$$\mathbb{E}[(x_{n+k} - \hat{x}^t_{t+k})x_h] = 0, \quad h = 1, \ldots, n.$$

Finding $\phi_{n,1}, \ldots, \phi_{n,n}$ such that $\hat{x}^t_{t+k} = \sum_{j=1}^n \phi_{n,j} x_{n+1-j}$ is again equivalent to solving

$$\sum_{i=1}^{n} \phi_{n,i}\gamma(i-j) = \gamma(j), \qquad j = 1, \ldots, n,$$

or in short

$$\boldsymbol{\Gamma_n \phi_n} = \boldsymbol{\gamma_n},$$

where $\boldsymbol{\Gamma_n} = [\gamma(i-j)]_{i,j=1,\ldots,n}$, $\boldsymbol{\gamma(n)} = (\gamma(1), \ldots, \gamma(n))$ with $\gamma(\cdot)$ the autocorrelation function of the model and $\boldsymbol{\phi_n} = (\phi_{n,1}, \ldots, \phi_{n,n})$

[1]The $AICc$ in its general form states that for $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_q)$,

$$AICc(\boldsymbol{\theta}, \boldsymbol{\phi}) = -2lnL\left(\boldsymbol{\theta}, \boldsymbol{\phi}, \frac{S(\boldsymbol{\theta}, \boldsymbol{\phi}, \sigma^2)}{n}\right) + \frac{2(p+q+1)n}{n-p-q-2},$$

where

$$L\left(\boldsymbol{\theta}, \boldsymbol{\phi}, \sigma^2\right) = (2\pi\sigma^2)^{-\frac{n}{2}}(r_0 \cdots r_{n-1})^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\sigma^{-2}\sum_{j=1}^{n}\frac{(x_j - \hat{x}_j)^2}{r_{j-1}}\right]$$

is the likelihood of the data under the Gaussian ARMA model with parameters $\left(\boldsymbol{\theta}, \boldsymbol{\phi}, \sigma^2\right)$ and $S(\boldsymbol{\theta}, \boldsymbol{\phi}, \sigma^2) = \sum_{j=1}^n \frac{(x_j - \hat{x}_j)^2}{r_{j-1}}$ is the residual sum of squares, $r_i = \frac{E(x_{i+1} - \hat{x}_{i+1})^2}{\sigma^2}$.

The model to be selected is the model with parameters $(p, q, \boldsymbol{\theta}, \boldsymbol{\phi})$ which minimize the $AICc$.

Instead of solving for $\boldsymbol{\phi_n} = \boldsymbol{\Gamma_n^{-1}\gamma_n}$ by inversion of the matrix $\boldsymbol{\Gamma_n}$ the coefficients can be solved using the *Innovations Algorithm*[2]. In the case of an $AR(1)$ model it simplifies drastically so that the prediction $\hat{x}^t_{t+k}$ becomes

$$\hat{x}^t_{t+k} = (\phi_1)^k x_t.$$

In the case of an $ARIMA(1,1,0)$ model, the Innovations Algorithm produces

$$\hat{x}^t_{t+k} = (\phi_1)^k(x_t - x_{t-1}) + (\phi_1)^{k-1}(x_t - x_{t-1}) + \cdots + (\phi_1)(x_t - x_{t-1}) + x_t.$$

Lastly, to determine the optimal length of the rolling training period, models with training periods of 12, 24, 36 and 48 months were fitted to the T-Bill rates from 1975 to year-end 1981. Forecasts were then compared using the root-mean-squared error. The autoregressive model performed best with a training period of 36 months and consequently this was the chosen length for the entire forecasting period from 1982 to 1996. This forecast was denoted by $AR_{36}(1)$. The R code used for fitting the autoregressive model, testing the best length for the training period and make predictions is printed in section D.1 (R Development Core Team, 2010).

4.1.4. Forward rate forecast

The *forward rate forecast* uses a forward rate for its predictions. A *forward rate* is an interest rate implied by the yield curve. The interest rates for all government securities with different maturities[3] are summarized by the *yield curve*. The yield curve displays

[2]The Innovations Algorithm for an $ARIMA(p,d,q)$ process solves the projection P_n of x_{t+k} onto $\hat{x}^t_{t+k} \in span(x_1, \ldots, x_n)$ via

$$P_t x^t_{t+k} = \sum_{j=1}^{p+d} \phi^*_j P_t x_{t+k-j} + \sum_{j=k}^{q} \theta_{t+k-1,j}(x_{t+k-j} - \hat{x}^t_{t+k-j}),$$

where ϕ^*_j are the coefficients of $\phi^*(z) = (1-z)^d\phi(z) = 1 - \phi^*_1 z - \cdots - \phi^*_{p+d} z^{p+d}$, and $\theta_{t+k-1,k}, \ldots \theta_{n+k-1,q}$ can be computed recursively from the equations

$$\nu_0 = \kappa(1,1),$$

$$\theta_{t+k-1,n+k-1-i} = \nu_i^{-1}\left(\kappa(t+k, i+1) - \sum_{j=0}^{k-1} \theta_{i,i-j}\theta_{t+k-1,t+k-1-i}\nu_j\right), \quad 0 \le i \le t+k-1,$$

and

$$\nu_{t+k-1} = \kappa(t+k, t+k) - \sum_{j=0}^{n-1} \theta^2_{t+k-1,t+k-1-j}\nu_j,$$

where $\kappa(i,j) = \mathbb{E}[x_i x_j]$.

[3]Maturities include 1-month, 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, 10-year, 20-year and 30-year government securities.

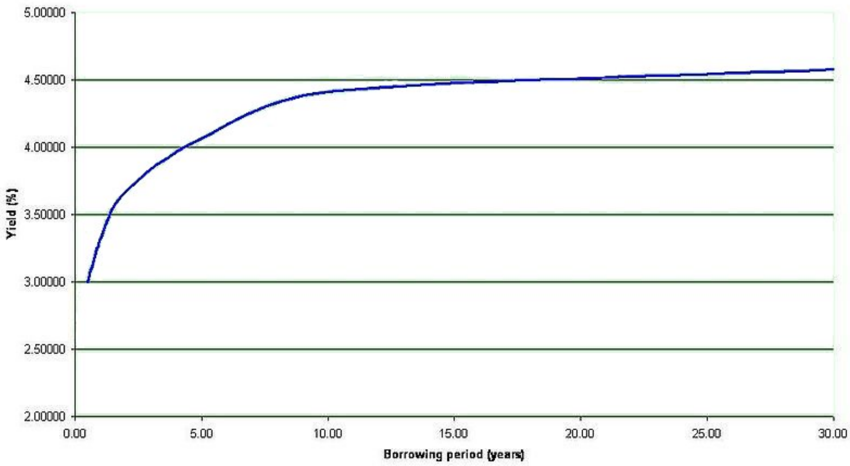


Figure 4.3.: Sample of a yield curve curve taken from February 9, 2005 (Treasury, 2010a)

the relationship between interest rates and time to maturity of Treasury bills as displayed in Figure 4.3.[4]

A forward rate then is an interest rate that is implied by the interest rates of two T-Bills of different maturities. If the interest rate of a bill with maturity m is known as well as the interest rate of a bill with maturity $m+n$ then this implies an interest rate of a bill issued at time m with maturity n (Gurkaynak *et al.*, 2007). The underlying assumption is again that of arbitrage-free pricing which assumes that equivalent investments over the same time period are perfect substitutes regardless of their underlying maturities (Hull, 2006). Consequently, arbitrage opportunities should not exist. The return of an investment on a bill with an annualized interest rate of i^0_{m+n} must be the same as the return of an investment in a bill with interest rate i^0_m followed by an investment in a bill issued at time m with interest rate i^m_{m+n}, that is,

$$(1 + i^t_{m+n})^{m+n} = (1 + i^t_m)^m (1 + i^{t+m}_{m+n})^n.$$

The forward rate can thus be obtained by solving for i^{t+m}_{m+n} (Fabozzi, 2005):[5]

$$i^{t+m}_{m+n} = \frac{1}{n}\left(\frac{1 + (m+n)i^t_{m+n}}{1 + mi^t_m} - 1\right).$$

[4]The upward slope can be explained by the demand for compensation by investors for longer exposure to risk and uncertainty, as is proposed by the *Liquidity Premium Theory*. However, this need not always be the case (Hull, 2006).

[5]The equation makes use of a very common approximation when calculating interest rates: $(1+i)^n \approx 1 + ni$ for small i.

In order to obtain forward rates of 3-month Treasury bills up to 9 months into the future it was necessary to reconstruct a yield curve that displays interest rates of maturities up to 12 months ahead. For the time period from 1982 to 1996, the Federal Reserve provides complete data for the 3-month, 6-month and 12-month interest rates of Treasury bills. The U.S. Treasury's yield curve is derived using a quasi-cubic Hermite spline function (Treasury, 2010b). Although, as noted on their website,

> "Treasury does not provide the computer formulation of our quasi-cubic hermite spline yield curve derivation program. However, we have found that most researchers have been able to reasonably match our results using alternative cubic spline formulas."

Indeed, several researchers use cubic spline interpolation for constructing the yield curve, as was done by Waggoner (1997), Fisher *et al.* (1995) and McCulloch (1975).

A *cubic spline* of a data set $\{i_1, \dots, i_n\}$ is a piecewise cubic polynomial real function $p : [a, b] \to \mathbb{R}$

$$p(z) = \begin{cases} p_1(z) & , \quad z_1 \le z \le z_2 \\ p_2(z) & , \quad z_2 \le z \le z_3 \\ & \vdots \\ p_{n-1}(z) & , \quad z_{n-1} \le x \le z_n \end{cases}$$

on an interval $[a, b]$ and subintervals $[z_j, z_{j+1}]$, $j = 1, \dots, n-1$, $a = z_1 < z_2 < \dots < z_{n-1} < z_n = b$, satisfying

1. each p_j is a cubic polynomial passing through its respective end points,

$$\begin{aligned} p_j(z_j) &= i_j \\ p_j(z_{j+1}) &= i_{j+1} \quad , \qquad j = 1, \dots, n-1 \end{aligned}$$

2. the first and second derivatives of adjoining polynomials p_j and p_{j+1} match at their shared end points,

$$\begin{aligned} \frac{d}{dz} p_j(z)|_{z=z_{j+1}} &= \frac{d}{dz} p_{j+1}(z)|_{z=z_{j+1}} \\ \frac{d^2}{dz^2} p_j(z)|_{z=z_{j+1}} &= \frac{d^2}{dz^2} p_{j+1}(z)|_{z=z_{j+1}} \quad , \qquad j = 1, \dots, n-2 \end{aligned}$$

3. the second derivatives are equal to zero at the end points of the interval,

$$\frac{d^2}{dz^2} p_1(z)|_{z=z_1} = 0 = \frac{d^2}{dz^2} p_{n-1}(z)|_{z=z_n}$$

A cubic spline then is derived from a system of $4(n-1)$ linear equations in $4(n-1)$ unknown variables. The set of 3-month, 6-month and 12-month T-Bills can then be used to calculate the forward rates i_{t+k+3}^{t+k}, $k = 1, \dots 9$ for every $t = 0, \dots 179$.

The forward rate forecast then issues the forecast

$$\hat{\boldsymbol{x}}^t = (i_{t+4}^{t+1}, i_{t+5}^{t+2}, \dots, i_{t+12}^{t+9})'.$$

The underlying rationale of the forward rate forecast is that all risk-free rates must match those implied by the yield curve in order to disallow arbitrage opportunities. Using a forward rate as a prediction for T-Bill rates therefore assumes no change in the current yield curve structure. Inaccuracies of the forecast are therefore due to changes of the yield curve.

4.1.5. Other forecasts

Leitch and Tanner (1991) use two additional forecasts for comparison. First, they have access to a "Professional forecast" which was published monthly by an institution called Commonwealth Research Group, starting in December 1981. Unfortunately, the author of this paper does not have access to those forecasts. Second, Leitch and Tanner use a survey forecast which was published together with the previously mentioned Professional forecast. It consists of consensus predictions by various experts from the financial sector. Although this data is unavailable as well, a very similar publication exists called the "Survey of Professional Forecasters" provided by the Federal Reserve Bank of Philadelphia (FRBP, 2010). It is freely available and also includes forecasts of T-Bill rates dating as far back as 1981. However, forecasts are done only quarterly and therefore do not allow for full comparison between all the aforementioned forecasts. For these reasons, this paper only makes use of the naive no-change, constant rate, autoregressive and forward rate forecasts.

4.2. Probabilistic forecasts

If X is a random variable on a probability space (Ω, Σ, P), a *probabilistic forecast* for X is a probability distribution F on Ω. A probabilistic forecast can be derived from a set of point forecasts by looking at the errors of the past m predictions by the point forecast. This probabilistic forecast shall be called the *error-corrected probabilistic* (*ecp*) forecast.

If a prediction with a k-month horizon is desired, the set of the previous m point predictions made with a k-month horizon is required. Let $Err_{k,m}^t$ be the set of the m preceding

k-month prediction errors of the point forecast $\hat{\boldsymbol{x}}^t$:

$$Err^t_{k,m} := \{x_{t-i+1} - \hat{x}^{t-i-k+1}_{t-i+1} | i = 1, \ldots, m\}.$$

Then define $PErr^t_{k,m}$ by adding the prediction $\hat{x}^t_{t+k}$ to each point in $Err_{\hat{x}^{t-m}_{t+k-m}}$:

$$PErr^t_{k,m} := \{\hat{x}^t_{t+k} + (x_{t-i+1} - \hat{x}^{t-i-k+1}_{t-i+1}) | i = 1, \ldots, m\}.$$

A probabilistic forecast $\hat{\overline{\boldsymbol{x}}}^t$ is then obtained via probability mass functions $P_{\hat{\overline{\boldsymbol{x}}}^t_{t+k}}(x)$ which assign a mass of $\frac{1}{m}$ to the points $\hat{x}^t_{t+k} + (x_{t-i+1} - \hat{x}^{t-i-k+1}_{t-i+1})$ that lie scattered around the predicted value of the point forecast $\hat{x}^t_{t+k}$. The probability mass function of $\hat{\overline{x}}^t_{t+k}$ associated with prediction $\hat{x}^t_{t+k}$ for every forecast horizon k is

$$P_{\hat{\overline{x}}^t_{t+k}}(\overline{x} = x) = \begin{cases} \frac{1}{m} & \text{if } x \in PErr^t_{k,m} \\ 0 & \text{otherwise} \end{cases}$$

Essentially the error-corrected probabilistic forecast "learns" from past prediction errors and "corrects" its predictions by subtracting each error from the point forecast. Every such "corrected" point forecast is then assigned an equal probability mass to obtain a discrete probability forecast.

Creating the ecp forecasts is done using R and the code is printed in section D.2.

5. Evaluation

5.1. Scoring functions

To evaluate the performance of a point forecast for a real-valued random variable, a scoring function can be put into practice. A *scoring function* is a mapping $S : D \times D \to [0, \infty)$ where $D \subseteq \mathbb{R}^n$ is the domain of both the predictions and the realized observations of the random variable in question. It allows a comparison of forecasts for a given set of predictions and observations. It is said to be *negatively oriented* if a forecast with a small score $S(x, y)$ is to be preferred over one with a larger score, and *positively oriented* otherwise. To evaluate a set of n point forecasts a *summary measure* is used by averaging the score function of each prediction $\hat{x}$ and respective observation x: $\frac{1}{n}\sum_{i=1}^{n} S(\hat{x}_i, x_i)$. The following statistical summary measures are used in this study.

The *average directional accuracy* (ADA) is a summary measure that rewards a correct prediction of the direction of the movement of the futures prices from one month to the following month. The indicator function $\mathbf{1}\{y_{t+k}^{t+1} \geq y_{t+k}^{t}\}$ is used to indicate a rise in futures prices. If the forecaster simultaneously predicts a rise in futures prices, expressed through $\mathbf{1}\{\hat{x}_{t+k}^{t} \geq y_{t+k}^{t}\}$, then he will receive a score of 1 and 0 otherwise. Together with an analogous scheme for predicting a fall in futures prices yields the summary measure

$$ADA(\hat{\boldsymbol{x}}, \boldsymbol{y}) = \frac{1}{9}\sum_{k=1}^{9} \frac{1}{n_k} \sum_{t=0}^{n_k-1} \left(\mathbf{1}\{y_{t+k}^{t+1} \geq y_{t+k}^{t}\}\mathbf{1}\{\hat{x}_{t+k}^{t} \geq y_{t+k}^{t}\} + \mathbf{1}\{y_{t+k}^{t+1} < y_{t+k}^{t}\}\mathbf{1}\{\hat{x}_{t+k}^{t} < y_{t+k}^{t}\}\right),$$

where n_k is the number of months for which predictions exist with a horizon of k months,

$$n_k = 181 - k \quad \text{for} \quad k = 1, \dots, 9.$$

The *mean absolute error* (MAE) is a summary measure which may be defined as

$$MAE(\hat{\boldsymbol{x}}, \boldsymbol{y}) = \frac{1}{9}\sum_{k=1}^{9} \frac{1}{n_k} \sum_{t=0}^{n_k-1} |\hat{x}_{t+k}^{t} - y_{t+k}^{t+1}|.$$

Furthermore there is the very commonly used *root-mean-squared error* (RMSE) summary measure:

Table 5.1.: Correlations between x_{t+k} and y_{t+k}^{t+1}, for every forecast horizon k, for a total of 1581 data pairs

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
x_{t+k} vs. y_{t+k}^{t+1}	1	0.982	0.958	0.935	0.909	0.879	0.845	0.820	0.803

$$RMSE(\hat{\boldsymbol{x}}, \boldsymbol{y}) = \frac{1}{9} \sum_{k=1}^{9} \sqrt{\frac{1}{n_k} \sum_{t=0}^{n_k-1} (\hat{x}_{t+k}^{t} - y_{t+k}^{t+1})^2}.$$

Note that instead of evaluating predictions $\hat{x}_{t+k}^t$ against the realizations x_{t+k}, the futures rate of the following month y_{t+k}^{t+1} is used instead. This is done in accordance with Leitch and Tanner (1993, p. 585):

> "Note that, instead of the actual realized interest rate, we have used only the next month's new futures-market forecasts for the realization in this table [see Table 5.6]. However, the results are not affected by doing the calculations this way, as the cash markets[1] give essentially identical results."

For all 1581 predictions that were made, the correlation between x_{t+k} and y_{t+k}^{t+1} was calculated for every forecast horizon to infer the interchangeability of the two. Results are displayed in Table 5.1. It can be seen that the two values do correlate strongly for short forecast horizons, as is expected, but not in a manner that one would expect identical results.

The *profit rules* are summary measures that determine the hypothetical profit that could be gained from a forecast:

Given that one is currently at month t, the current T-Bill price x_t and the futures prices for the following 9 months $y_{t+1}^t, y_{t+2}^t, \ldots, y_{t+9}^t$ are observed. If one agrees to buy a futures contract for month $t+k$, then one agrees that at month $t+k$ one will make a payment of y_{t+k}^t to buy a T-Bill which can then be sold for x_{t+k}. The resulting profit (or loss) will be $x_{t+k} - y_{t+k}^t$.

Vice versa, if one agrees to sell a futures contract for month $t+k$, then one agrees to deliver a T-Bill at month $t+k$ for which a payment of y_{t+k}^t will be received. If the T-Bill is not previously owned one will have to buy it first, the price for which will be x_{t+k}, resulting in profit (or loss) $y_{t+k}^t - x_{t+k}$.

On a fictious $1 million three-month Treasury bill that is held until maturity, each basis-point change in the interest rate changes the gross return by $2500 ($1 million x 0.01

[1]"Cash market" simply refers to the market for Treasury bills as opposed to the market for its futures contracts.

x 0.25 years). For all four profit rules a \$40 cost for every purchase/sale of a contract is assumed, which is an estimate of transaction costs (broker's fees) and lost interest on the money invested. The summary measures $\pi(\hat{\boldsymbol{x}}, \boldsymbol{y})$ of the profit rules then express the average annual profit[2]. Given a prediction for a T-Bill, four alternative strategies for buying and selling are considered, called *profit rules.*

Profit rule A instructs to buy a futures contract y^t_{t+k} for month $t+k$ if a rise in T-Bill prices is predicted, i.e. $\hat{x}^t_{t+k} > x_t$. Conversely, one sells a futures contract if $\hat{x}^t_{t+k} < x_t$ is predicted. If the prediction is for no change, $\hat{x}^t_{t+k} = x_t$, as is the case with the no-change forecast, no action is taken. The implicit assumption in using this profit rule is that the market does not expect interest rates to change.

$$\begin{aligned}\pi_A(\hat{\boldsymbol{x}}, \boldsymbol{y}) = \sum_{k=1}^{9} \frac{1}{m} \sum_{t=0}^{n_k} \Big(& \mathbf{1}\{\hat{x}^t_{t+k} > x_t\} \quad \underbrace{(x_{t+k} - y^t_{t+k})}_{\text{profit/loss from buying futures}} \\ & + \mathbf{1}\{\hat{x}^t_{t+k} < x_t\} \quad \underbrace{(y^t_{t+k} - x_{t+k})}_{\text{profit/loss from selling futures}} \\ & - \mathbf{1}\{\hat{x}^t_{t+k} \neq x_t\}(\$40)\Big),\end{aligned}$$

where m is the duration of the forecasting period in years.

Profit rule B states to buy a futures contract y^t_{t+k} for month $t+k$ if it is predicted that the rate will lie above the futures market rate, i.e. $\hat{x}^t_{t+k} > y^t_{t+k}$. Conversely, one sells a futures contract if $\hat{x}^t_{t+k} < y^t_{t+k}$ is predicted. In contrast to profit rule A, the assumption that the market expects no change of rates is dropped:

$$\begin{aligned}\pi_B(\hat{\boldsymbol{x}}, \boldsymbol{y}) = \sum_{k=1}^{9} \frac{1}{m} \sum_{t=0}^{n_k} \Big(& \mathbf{1}\{\hat{x}^t_{t+k} > y^t_{t+k}\} \quad \underbrace{(x_{t+k} - y^t_{t+k})}_{\text{profit/loss from buying futures}} \\ & + \mathbf{1}\{\hat{x}^t_{t+k} < y^t_{t+k}\} \quad \underbrace{(y^t_{t+k} - x_{t+k})}_{\text{profit/loss from selling futures}} \\ & - \mathbf{1}\{\hat{x}^t_{t+k} \neq y^t_{t+k}\}(\$40)\Big).\end{aligned}$$

Profit rule C is the same as B with the additional requirement that a change must be predicted, $\hat{x}^t_{t+k} \neq x_t$, for any action to be taken. If $\hat{x}^t_{t+k} = x_t$ is predicted, no futures contract is bought or sold:

[2]Strictly speaking, the profit rules used are not an annual average in the conventional sense where one would average each year's profit. Instead, profits are averaged over every forecast horizon and then summed up in order to stay consistent with the methods of the other summary measures.

$$\begin{aligned}\pi_C(\hat{\boldsymbol{x}}, \boldsymbol{y}) = \sum_{k=1}^{9} \frac{1}{m} \sum_{t=0}^{n_k} \Big(\mathbf{1} \underbrace{(\hat{x}_{t+k}^t \neq x_t)}_{\text{take a position}} \big[\mathbf{1}\{\hat{x}_{t+k}^t > y_{t+k}^t\} \quad & \underbrace{(x_{t+k} - y_{t+k}^t)}_{\text{profit/loss from buying futures}} \\ + \mathbf{1}\{\hat{x}_{t+k}^t < y_{t+k}^t\} \quad & \underbrace{(y_{t+k}^t - x_{t+k})}_{\text{profit/loss from selling futures}} \\ - \mathbf{1}\{\hat{x}_{t+k}^t \neq y_{t+k}^t\}(\$40) \big] \Big). &\end{aligned}$$

Profit rule D orders to buy a futures contract if a fall in rates is predicted while the futures market rate lies above the current rate, i.e. $\hat{x}_{t+k}^t < x_t$ and $y_{t+k}^t > x_t$. Conversely, one sells a contract if a rise is predicted while the futures price lies below the current rate: $\hat{x}_{t+k}^t > x_t$ and $y_{t+k}^t < x_t$. On occasions where both the forecast and the futures price lie above the current rate, i.e. $\hat{x}_{t+k}^t > x_t$ and $y_{t+k}^t > x_t$, no action is taken. Likewise if both lie below the current rate:

$$\begin{aligned}\pi_D(\hat{\boldsymbol{x}}, \boldsymbol{y}) = \sum_{k=1}^{9} \frac{1}{m} \sum_{t=0}^{n_k} \Big(\mathbf{1}\{\hat{x}_{t+k}^t > x_t\}\mathbf{1}\{y_{t+k}^t < x_t\} \quad & \underbrace{(x_{t+k} - y_{t+k}^t)}_{\text{profit/loss from buying futures}} \\ + \mathbf{1}\{\hat{x}_{t+k}^t < x_t\}\mathbf{1}\{y_{t+k}^t > x_t\} \quad & \underbrace{(y_{t+k}^t - x_{t+k})}_{\text{profit/loss from selling futures}} \\ - \mathbf{1}\{\hat{x}_{t+k}^t \neq y_{t+k}^t\}(\$40) \Big). &\end{aligned}$$

A final note on the summary measures: when averaging the performances of the predictions under each scoring function, slight variations of the interpretation of "average" are possible. The summary measures employed all firstly average a forecast over every forecast horizon before averaging the performances per horizon, as is expressed by $\frac{1}{9}\sum_{k=1}^{9} \frac{1}{n_k} \sum_{t=0}^{n_k-1}(\dots)$. Alternatively, one could average over all predictions, giving each an equal weighting in the form of $\frac{1}{n}\sum_{k=1}^{n}(\dots)$ where n is the total number of predictions. For this study, however, the resulting differences between interpretations is negligible with regard to the final results.

5.2. Bayes rules

Arriving at a point prediction from a probabilistic forecast requires that the forecaster is instructed to provide a functional of his predictive distribution (e.g. the mean functional) and that the scoring function for evaluation is consistent[3] for the functional; alternatively, the forecaster can be provided the scoring function, in which case he is able to make an *optimal point forecast* $\hat{x}^t$ by using the Bayes rule

$$\hat{x}^t = \arg\max_x \mathbb{E}_F[S(x, Y)]$$

for a positively oriented scoring function S (using arg min instead for a negatively oriented score function), where F is the predictive distribution for the random variable Y. The argument x which maximizes $\mathbb{E}_F S(x, Y)$ is dependent on S. It is for this reason that a "best" forecast is meaningless in the context of point forecasts when the scoring function used for evaluating the forecast is unknown (Gneiting, 2010). Considerations have to be made which point forecast to deliver for each of the scoring functions in use.

For the average directional accuracy, the Bayes rule yields

$$\begin{aligned}\mathbb{E}_F[ADA(\hat{\boldsymbol{x}}, \boldsymbol{y})] = \Big(&\mathbb{E}_F[\mathbf{1}\{y^{t+1}_{t+k} \geq y^t_{t+k}\}]\mathbf{1}\{\hat{x}^t_{t+k} \geq y^t_{t+k}\} \\ &+ \mathbb{E}_F[\mathbf{1}\{y^{t+1}_{t+k} < y^t_{t+k}\}]\mathbf{1}\{\hat{x}^t_{t+k} < y^t_{t+k}\}\Big) \\ &= \left([1 - F(y^t_{t+k})]\mathbf{1}\{\hat{x}^t_{t+k} \geq y^t_{t+k}\} + [F(y^t_{t+k})]\mathbf{1}\{\hat{x}^t_{t+k} < y^t_{t+k}\}\right).\end{aligned}$$

If

$$\begin{aligned}1 - F(y^t_{t+k}) &\geq F(y^t_{t+k}) \\ \Leftrightarrow median(F) &\geq y^t_{t+k},\end{aligned}$$

the expected value of the scoring function is maximized for any $\hat{x}^t_{t+k} \geq y^t_{t+k}$ vis-a-vis $\hat{x}^t_{t+k} < y^t_{t+k}$ if $median(F) < y^t_{t+k}$.

For the mean absolute error, the median μ_{med} of a random variable X minimizes the score, as can be demonstrated briefly. If P is a probability measure, $a \in \mathbb{R}$ and without loss of generality $a < \mu_{\text{med}}$, then

$$\begin{aligned}|X - a| - |X - \mu_{\text{med}}| = (a - \mu_{\text{med}})\mathbf{1}\{X < a\} &+ (2X - a - \mu_{\text{med}})\mathbf{1}\{a < X < \mu_{\text{med}}\} \\ &+ (\mu_{\text{med}} - a)\mathbf{1}\{X \geq \mu_{\text{med}}\}\end{aligned}$$

$$\begin{aligned}\Rightarrow \mathbb{E}_P[|X - a| - |X - \mu_{\text{med}}|] &= \underbrace{(\mu_{\text{med}} - a)}_{\geq 0}\underbrace{(1 - 2(1 - \mathrm{P}(X \geq \mu_{\text{med}})))}_{\geq 0} \\ &+ 2\,\mathbb{E}_P[\underbrace{(X - a)\mathbf{1}\{a < X < \mu_{\text{med}}\}}_{\geq 0}] \\ &\geq 0.\end{aligned}$$

[3]A scoring function S is *consistent* for a functional T relative to the class $\mathcal{F}$ of probability measures on the observation domain D if $\mathbb{E}_F[S(t, Y)] \leq \mathbb{E}_F[S(x, Y)]$ for all probability measures $F \in \mathcal{F}$, all $t \in T(F)$ and all $x \in D$.

Similarly, it can be shown that the mean μ of a random variable X minimizes quadratic scoring functions. If $\mu_2 = \mathbb{E}_P[X^2] - \mu^2$, then

$$\begin{aligned}\mathbb{E}_P[(X-a)^2] =& \mathbb{E}_P[(X-\mu+\mu-a)^2] \\ =&\mu_2 + (a-\mu)^2\end{aligned}$$

so that the score is minimal for $a = \mu$, where $\mu = \mathbb{E}_P[X]$.

As for the profit rules, a Bayes rule is obtained via

$$\begin{aligned}\mathbb{E}_F[\pi_A(\hat{\boldsymbol{x}}, \boldsymbol{y})] = & (\mathbf{1}\{\hat{x}^t_{t+k} > x_t\}(\mu - y^t_{t+k}) \\ & + \mathbf{1}\{\hat{x}^t_{t+k} < x_t\}(y^t_{t+k} - \mu) \\ & - \mathbf{1}\{\hat{x}^t_{t+k} \neq x_t\}\{\$40\}),\end{aligned}$$

indicating that any $\hat{x}^t_{t+k} > x_t$ is optimal if $y^t_{t+k} - \mu > \mu - y^t_{t+k}$ and vice versa. [4] Profit rule B has an analogous Bayes rule, the only difference being that any $\hat{x}^t_{t+k} > y^t_{t+k}$ becomes optimal if $y^t_{t+k} - \mu > \mu - y^t_{t+k}$ and vice versa. Profit rule C in turn has the same Bayes rule as Profit rule B with the only addition that $\hat{x}^t_{t+k} \overset{!}{\neq} x_t$ in order to realize positive profit. For profit rule D, given $y^t_{t+k} - \mu > 0$, it's also necessary that $y^t_{t+k} < x_t$ in order for positive profits to be attainable. If the latter condition is not met, any $\hat{x}^t_{t+k}$ is optimal since profit will be 0 anyway. Otherwise, any $\hat{x}^t_{t+k} > x_t$ is a Bayes rule, and vice versa for a change of signs.

The Bayes rules are summarized in Table 5.2

5.3. Results 1982 to 1996

The performance of each forecast from the studied period from 1982 to 1996 can be seen in Table 5.3. The profit rules are expressed in thousands and always rounded to the nearest thousand.

It is apparent that the naive no-change forecast is the most competitive of the forecasts: it performs best of all the forecasts under 5 of the 7 scoring functions. It can be seen from the figures in Appendix A that this is due to the naive no-change forecast's strong performance over long forecast horizons. The overall strong performance of the naive no-change forecast is agreeable with the random walk assumption of financial markets. Under this assumption the naive no-change forecast is optimal.

[4] An additional necessary condition is that the potential profit outweigh the transaction costs, $|\mu - y^t_{t+k}| > \$40$ but this is always met with even the smallest possible change in interest rates of 0.01%

Table 5.2.: Bayes rules for the various scoring functions

Scoring function	Bayes rule	Condition
ADA	$\hat{x}^t_{t+k} \geq y^t_{t+k}$	$\text{median}(F) \geq y^t_{t+k}$
	$\hat{x}^t_{t+k} < y^t_{t+k}$	$\text{median}(F) < y^t_{t+k}$
MAE	$\hat{x}^t_{t+k} = \text{median}(F)$	
RMSE	$\hat{x}^t_{t+k} = \text{mean}(F)$	
Profit A	$\hat{x}^t_{t+k} > x_t$	$\mu - y^t_{t+k} > 0$
	$\hat{x}^t_{t+k} < x_t$	$y^t_{t+k} - \mu > 0$
Profit B	$\hat{x}^t_{t+k} > y^t_{t+k}$	$\mu - y^t_{t+k} > 0$
	$\hat{x}^t_{t+k} < y^t_{t+k}$	$y^t_{t+k} - \mu > 0$
Profit C	$\hat{x}^t_{t+k} > y^t_{t+k}$, $\hat{x}^t_{t+k} \neq x_t$	$\mu - y^t_{t+k} > 0$
	$\hat{x}^t_{t+k} < y^t_{t+k}$, $\hat{x}^t_{t+k} \neq x_t$	$y^t_{t+k} - \mu > 0$
Profit D	$\hat{x}^t_{t+k} \geq x_t$	$\mu - y^t_{t+k} > 0$ and $y^t_{t+k} < x_t$
	any $\hat{x}^t_{t+k}$	$\mu - y^t_{t+k} > 0$ and $y^t_{t+k} \geq x_t$
	$\hat{x}^t_{t+k} < x_t$	$y^t_{t+k} - \mu > 0$ and $y^t_{t+k} \geq x_t$
	any $\hat{x}^t_{t+k}$	$y^t_{t+k} - \mu > 0$ and $y^t_{t+k} < x_t$

The constant rate forecast performs worst under every scoring function. Figure A.1.2 as well as the according figures in Appendix B give evidence to its dismal performance over long forecast horizons. This is particularly true during periods of large price changes as was the case in 1982 and 1988 where the constant rate forecast is most susceptible to large prediction errors.

Although the autoregressive forecast fails to outperform the naive no-change forecast, note its strong performance under short forecast horizons of 1 to 3 months as demonstrated by the figures in Appendix B.

The forward rate forecast is interesting in that it performs well under the profit measures while doing relatively badly when using the standard measures. Upon closer inspection, it becomes noticeable from the figures in Appendix B that the forward rate forecast appears to have an erratic performance when comparing forecast horizons as opposed to having a clear trend as one would expect. Notice, however, that performance generally is best for a 3-month and 6-month horizon which also happens to knots on the yield curve for which the interest rates are provided. The bad performance in between these knots can therefore be traced back to the construction method of the yield curve which does not seem to provide a good forward rate prediction. In particular, the low performance over 1- and 2-month horizons is conspicuous, which gives evidence that the yield curve may not be modeled well by a cubic polynomial over the 0- to 3-month term structure. Since the yield curve must necessarily go through the knots 0 at $t = 0$ and the spot rate of a 3-month T-Bill at $t = 3$, a cubic fitting may not be appropriate to model such short-term interest rates.

Table 5.3.: Forecast evaluation results, 1982 to 1996

	ADA	MAE	RMSE	Profit A	Profit B	Profit C	Profit D
Naive no-change	0.316	0.482	0.645	\$0	-\$106	\$0	\$0
Constant rate	0.601	1.186	1.904	-\$155	-\$186	-\$186	-\$130
$\text{AR}_{36}(1)$	0.558	0.652	0.910	-\$115	-\$126	-\$125	-\$110
Forward rate	0.416	1.141	1.440	\$2	-\$57	-\$57	-\$52
Ecp naive no-change	0.316	0.779	0.966	\$53	-\$10	-\$9	-\$27
Ecp constant rate	0.372	0.774	1.112	\$36	-\$36	-\$36	-\$35
Ecp $\text{AR}_{36}(1)$	0.534	0.873	1.131	-\$96	-\$120	-\$120	-\$101
Ecp forward rate	0.356	0.571	0.803	\$5	-\$73	-\$72	-\$51

The error-corrected probabilistic forecasts provide an improvement over their counterpart forecasts in some cases but not all. Most noticeably, the ecp forecasts perform much better under the profit rules and likewise under the average directional accuracy. This holds particularly for the constant rate and autoregressive forecast which will predict a price change in the same direction over every forecast horizon. This apparent weakness in prediction is corrected by the ecp forecasts and as a result, profits are increased for these two forecasts.

It may also come as a surprise that most of the forecasts produce negative annual profit. This fact cannot be explained by the accuracy (or rather, the lack thereof) of forecasts alone as the prevailing futures rates determine profits and losses. Even if a forecast correctly predicts a fall in T-Bill rates, if the corresponding futures rates is below the predicted value, according to profit rules A, B and C one would still opt to buy a futures contract. If T-Bill rates then fall below the rate of the futures contract, the contract results in a loss even though the correct price movement was predicted. As prices of financial markets are said to reflect all information available to investors by the efficient-market hypothesis, the generally negative profits give testimony that such information includes more than just knowledge of past rates, whereas the forecasts used rely solely on past and current rates.

Table 5.4 shows the correlations between each scoring function. Every forecast horizon is treated separately, meaning for every forecast, predictions over the same horizon were grouped together and then evaluated. The correlations between scoring functions uses the standard Pearson correlation

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}.$$

Table 5.4.: Correlations between the various scoring functions with each forecast horizon treated separately, 1982 to 1996

	ADA	MAE	RMSE	Profit A	Profit B	Profit C	Profit D
ADA	-	0.409	0.452	-0.753	-0.569	-0.654	-0.691
MAE	-	-	0.974	-0.366	-0.406	-0.479	-0.586
RMSE	-	-	-	-0.454	-0.484	-0.562	-0.654
Profit A	-	-	-	-	0.821	0.857	0.852
Profit B	-	-	-	-	-	0.892	0.797
Profit C	-	-	-	-	-	-	0.943

5.4. Results 1982 to 1987

Additionally, the results were cropped to the time window used by Leitch and Tanner (1991) to allow for better comparisons. Table 5.5 shows the corresponding results of this paper while Table 5.6 displays the results of Leitch and Tanner. Tables 5.7 and 5.8 again show the correlations found from each study.

Comparing Table 5.3 to the cropped results of Table 5.5, the discussed observations retain their validity. What is noticeable is the overall deteriorating of performances under every scoring rule. This is due to the nature of the data as the period from 1982 to 1983 saw an exceptionally large change in rates which the forecasts failed to capture.

The results of Leitch and Tanner (1991) from Table 5.6 have similar tendencies but obviously differ. A detailed discussion of possible reasons for the unexpected discrepancy is given in chapter 6. The conclusions drawn from the results run counter to those offered by Leitch and Tanner (1991): while correlations between profits and the standard measures may not be substantial, they are significantly higher than was observed in the Leitch and Tanner study.

Table 5.5.: Forecast evaluation results, 1982 to 1987

	ADA	MAE	RMSE	Profit A	Profit B	Profit C	Profit D
Naive no-change	0.299	0.574	0.747	$0	-$209	$0	$0
Constant rate	0.600	1.671	2.650	-$224	-$253	-$253	-$216
$AR_{36}(1)$	0.580	0.799	1.125	-$216	-$234	-$232	-$211
Forward rate	0.420	1.481	1.813	-$9	-$123	-$123	-$109
Ecp naive no-change	0.352	0.887	1.034	$33	-$89	-$88	-$87
Ecp constant rate	0.386	1.031	1.344	$73	-$29	-$28	-$68
Ecp $AR_{36}(1)$	0.585	1.089	1.415	-$242	-$251	-$250	-$225
Ecp forward rate	0.343	0.772	1.079	$72	-$92	-$92	-$68

Table 5.6.: Forecast evaluation results by Leitch and Tanner (1991), Table 1, 1982 to 1987

	ADA	MAE	RMSE	Profit A
Naive no-change	0.379	0.410	0.530	$0
Constant rate	0.466	2.013	2.514	-$674
AR(2)	0.479	0.739	0.902	-$928
Forward rate	0.437	0.656	0.848	-$3,050

Table 5.7.: Correlations between the various scoring functions with each forecast horizon treated separately, 1982 to 1987

	ADA	MAE	RMSE	Profit A	Profit B	Profit C	Profit D
ADA	-	0.467	0.516	-0.789	-0.483	-0.689	-0.725
MAE	-	-	0.974	-0.496	-0.307	-0.493	-0.581
RMSE	-	-	-	-0.566	-0.374	-0.558	-0.634

Table 5.8.: Correlations between the various scoring functions by Leitch and Tanner (1991), Table 4, with each forecast horizon treated separately, 1982 to 1987

	ADA	MAE	RMSE	Profit A	Profit B	Profit C	Profit D
ADA	-	0.012	0.024	0.441	0.819	0.619	0.572
MAE	-	-	0.996	-0.095	0.074	0.100	0.212
RMSE	-	-	-	-0.101	0.096	0.123	0.216

6. Conclusion

The purpose of this study was to perform forecasts of historical interest rates and analyze their performance under different scoring functions. The procedure was largely a replication of a 1991 study by Leitch and Tanner to assess their findings with since then newly available data.

Additionally, probabilistic forecasting was introduced in the form of error-corrected probabilistic forecasts derived from point forecasts. Probabilistic forecasts have the advantage that rather than limiting a future event to a single outcome, a forecast in form of a distribution shows a forecaster's full set of beliefs. Point forecasts may still be acquired from a probabilistic distribution where it suits practicality; however, it is imperative that they minimize expectation under the scoring function that is used to evaluate the prediction.

To conduct the forecasts, 3-month U.S. Treasury bill rates were gathered along with corresponding futures rates for the period from 1982 to year-end 1996.

At the end of each month, predictions were made for the T-Bill rates for the following 9 months. The forecasting techniques included a naive no-change forecast, a constant rate forecast, an autoregressive forecast and a forward rate forecast. Additionally, an *ecp* forecast was derived from each technique.

Two types of scoring functions were then employed to evaluate forecasts: standard measures and profit measures. Standard measures included the average directional accuracy, the mean average error and the root-mean-squared error. The profit measures consisted of a set of 4 profit rules that evaluated hypothetical profits from each forecast by virtually buying and selling T-Bill futures contracts at the prevailing futures rates.

Comparing the presented results to those of Leitch and Tanner (1991), this study partially diverges from theirs: in particular, Leitch and Tanner note the low correlation between the standard performance measures and profit measures. Such correlations were found to be significantly higher in this study, both over the 1982-1987 and 1982-1996 periods. In trying to explain this discrepancy, one notices that already the performances of forecasts under each performance measure differ when they should be the same. There are three identifiable sources of error for this.

Firstly, it is not precisely clear to the author exactly which rates of the 3-month Treasury bills were used. The Federal Reserve Statistical Release (FRB, 2010) offer rates both from the primary and secondary market. Furthermore, rates are available either day by day or as monthly averages. However, regardless of which rates are used, none delivered results closer to those offered by Leitch and Tanner. Each possibility was tested and the differences were only slim.

Secondly, some of the forecasting methods were difficult to reproduce: from the description provided it was not exactly clear how the constant rate forecast was done. An alternative interpretation to the one used in section 4.1.2 is that changes from the month previous to the current month are observed and the difference linearly extrapolated for each forecast horizon, i.e.

$$\hat{\boldsymbol{x}}^t = (x_t + (x_t - x_{t-1}), x_t + 2(x_t - x_{t-1}), \ldots, x_t + 9(x_t - x_{t-1}))'.$$

This method too was undertaken and yielded results no closer than the chosen method. Concerning the autoregressive forecast, the authors do not explain how the $ARIMA$ model was fitted to the training period from 1975 to 1981, only the resulting coefficients are presented. Using their model to make predictions also did not lead to the expected results, so a new AR fitting was done from scratch. The forward rate was difficult to reproduce because a reconstruction of the yield curve had to be done for which no definitive method exists, as explained in section 4.1.4. However, most surprisingly of all, the unambiguous naive no-change forecast did not match the results either.

This leads to a third possible error: the performance measures could have been applied differently than was done in this paper. Each was duplicated to the best of the author's ability since no formulas were provided. The average directional accuracy described herein adheres to the description by Leitch and Tanner (1991, p. 584) and measures

> "the percentage of interest rate changes in the futures market that were accurately forecast by each technique over the one-month observation interval until the new forecasts were available."

The mean average error and root-mean-squared error are common scoring functions with no ambiguity. Concerning these scoring functions, it was unclear, however, how predictions were averaged. As noted in section 5.1, different interpretations were undertaken and no meaningful differences were found. The profit rules are described in great detail, and the profit rules used here are simply an algebraic expression of the text description.

It remains a nuisance that the results could not be replicated. Nonetheless, the methods used in this study are very similar and have their own merits and justification, so it still is a noteworthy concluding remark that correlations between standard measures and profit measures were found to be more significant than was by Leitch and Tanner.

The benefits of applying an error-corrected probabilistic forecast to a point forecast are somewhat inconclusive. The *ecp* forecast generally outperformed its point forecast counterpart when performances were comparatively bad, as was the case for the constant rate forecast and the forward rate forecast. Performances actually worsened in general for comparatively good performances, such as the naive no-change and autoregressive forecast. The rational of applying an *ecp* forecast is to eliminate systematic errors that occur with forecasts that are susceptible to trends. This applied to the constant rate and forward rate forecast, and as expected, performances were generally improved.

What may strike the reader as remarkable is the fact that the most simplistic of the forecasting methods, namely the naive no-change forecast, is also one of the most competitive. While it may be the state of affairs that other, more sophisticated forecasts don't easily outperform the naive no-change forecast, one wouldn't necessarily expect them to do worse either. This result could be explained by the nature of the data: particularly the period of 1982 to 1985 saw extraordinarily high interest rates compared to the entire period since recording history began in 1934. This was mainly due to fiscal policy measures of the U.S. Federal reserve to control inflation at the time. With high fluctuations from month to month, forecasts that rely on earlier rates can be grossly misleading as is exemplarily seen from the constant rate forecast. Further insight is given by the figures in Appendix A: other forecasts perform better over short horizons but the naive no-change forecast does far better over long horizons. The strong performance of the naive no-change forecast is consistent with the random walk assumption of financial markets for which the naive no-change forecast is optimal.

Another conspicuous fact is the forward rate performing worst over horizons of 1 and 2 months compared to longer horizons when the opposite is expected. This can be explained by the construction technique of the yield curve: fitting a cubic spline polynomial to interpolate between the given points on the yield curve is problematic in that it does not model the yield curve well between the first and second knot. The first knot has to necessarily correspond to an interest rate of 0%. However, it seems unlikely that this first interval between the first and second knot is modeled well by a cubic polynomial. This is evidenced by the forecasts over a 1-month and 2-month horizon which rely on interpolated points between the first and second knot: they consistently over-estimate the interest rate and in doing so generally score worse than the predictions over the other horizons. Perhaps this issue is not important for most uses of yield curve interpolation with cubic splines because the time span covered by the first interval of this yield curve (90 days) is quite short compared to the entire span of the full yield curve (30 years). It is also noticeable how predictions are generally best for the 3 and 6 month horizons. Such predictions rely on knots from the yield curve that did not have to be constructed, thus perhaps displaying the limitations of cubic spline interpolation.

That every scoring function differs in how predictions are evaluated is not the contentious issue of this paper. The merits and appropriateness of a scoring function depend on the context and purpose of each forecast. Rather, it is interesting to note that the relationship, that is to say the correlation, between scoring functions does not appear to be the same as was established by Leitch and Tanner (1991). However, while the results here presented do not fully agree with those of the two authors, their findings have already found backing in the scientific community as is exhibited by 225 citations of their published paper.[1] Since the evaluation of forecasts is by no means a trivial matter, it seems vital that further investigation be undertaken to examine the relationship between standard statistical measures and ones expressing tangible measures such as profits.

[1]Citations performed via Google Scholar.

Bibliography

Abramson, Bruce, & Finizza, Anthony. 1995. Probabilistic forecasts from probabilistic models: a case study in the oil market. *International Journal of Forecasting*, **11**, 63–72.

Britton, Erik, Fisher, Paul, & Whitley, John. 1998. The Inflation Report projections: understanding the fan chart. *Bank of England Quarterly Bulletin*, **February**.

Brockwell, Peter J., & Davis, Richard A. 2006. *Time Series: Theory and Methods*. 2nd edn. Springer.

Cecchetti, Stephen G. 2006. *Money, Banking, and Financial Markets*. McGraw-Hill/Irwin.

CME. 2010. *Chicago Mercantile Exchange*. `http://www.cmegroup.com/market-data/datamine-historical-data/datamine.html`. [Online; accessed 19-May-2010].

Diebold, Francis X., & Li, Canlin. 2006. Forecasting the term structures of government bond yields. *Journal of Econometrics*, **130**, 337–364.

Fabozzi, Frank. 2005. *The Handbook of Fixed Income Securities*. McGraw-Hill.

Fisher, Mark, Nychka, Douglas, & Zervos, David. 1995. *Fitting the term structure of interest rates with smoothing splines*. Tech. rept.

FRB. 2010. *Federal Reserve Statistical Release H.15 - Historical Data*. `http://www.federalreserve.gov/boarddocs/hh/2004/july/testimony.html`. [Online; accessed 3-April-2010].

FRBP. 2010. *Federal Reserve Bank of Philadelphia*. `http://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/1`. [Online; accessed 28-May-2010].

Garbade, Kenneth D. 2008. Why the U.S. Treasury began auctioning Treasury bills in 1929. *Federal Reserve Bank of New York*, **14**(1).

Gneiting, Tilmann. 2010. Making and evaluating point forecasts. *arXiv:0912.0902v2*.

Granger, Clive W. J., & Pesaran, M. Hashem. 2000. Economic and statistical measures of forecast accuracy. *Journal of Forecasting*, **19**, 537–560.

Gurkaynak, Refet, Sack, Brian, & Wright, Jonathan. 2007. The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, **54**, 2291–2304.

Hull, John C. 2006. *Optionen, Futures und andere Derivate*. 6th edn. Pearson Studium.

Kwiatkowski, Denis, Phillips, Peter C.B., Schmidt, Peter, & Shin, Yongcheol. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, **54**, 159–178.

Labuszewski, John W., & Sturm, Frederick. *Unterstanding U.S. Treasury Futures*. `http://www.cmegroup.com/trading/interest-rates/files/Understanding_US_Treasury_Futures.pdf`. [Online; accessed 6-May-2010].

Leitch, Gordon, & Tanner, J. Ernest. 1991. Economic forecast evaluation: Profits versus the conventional error measures. *The American Economic Review*, **81**(3), 580–590.

Mankiw, N. Gregory. 1997. *Macroeconomics*. 3rd edn. Worth Publishers.

McCulloch, J. Huston. 1975. The tax-adjusted yield curve. *Journal of Finance*, **30**(1), 811–830.

Önkal, Dilek, & Muradoglu, Gülnur. 1994. Evaluating probabilistic forecasts of stock prices in a developing stock market. *European Journal of Operational Research*, **74**, 350–358.

Pesaran, M. Hashem, & Skouras, Spyros. 2002. *A Companion to Economic Forecasting*. Blackwell Publishers.

R Development Core Team. 2010. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Trapletti, Adrian, & Hornik, Kurt. 2011. *tseries: Time Series Analysis and Computational Finance*. R package version 0.10-25.

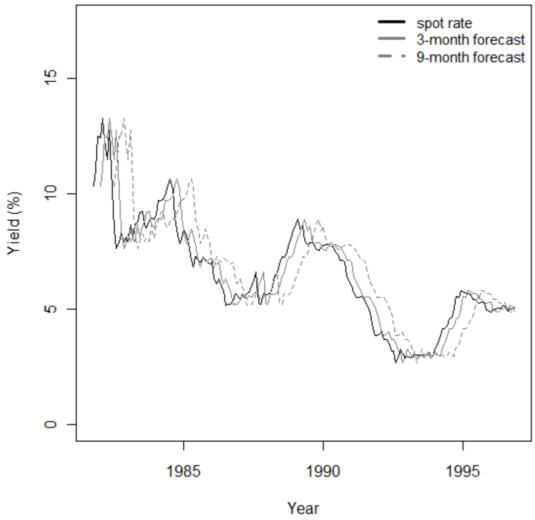
Treasury. 2010a. *Daily Treasury Yield Curve Rates*. `http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldYear&year=2005`. [Online; accessed 9-September-2010].

Treasury. 2010b. *Treasury Yield Curve Methodology*. `http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/yieldmethod.aspx`. [Online; accessed 9-September-2010].

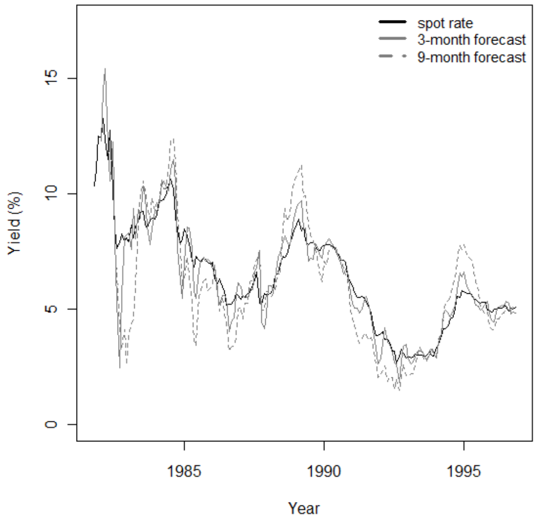
Waggoner, Daniel. 1997. Spline methods for extracting interest rate curves from coupon bond prices. *Federal Reserve Bank of Atlanta Working Paper*, **97**(10).

A. Forecast graphics

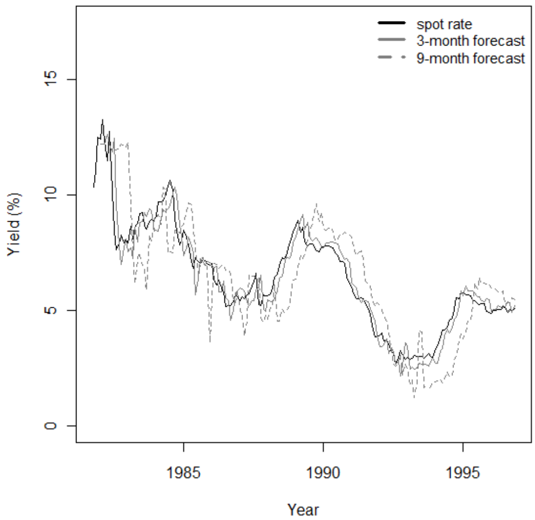
The following figures show the predictions made by every forecast over a 3-month and 9-month horizon. For comparison, the realized spot rates are included in each figure. The strength of the no-change forecast is visible from its 9-month forecast which is considerably closer to the realized spot rates than any of its competitors. However, one can also see that over the 3-month horizon, other forecasts such as the autoregressive forecast produce better predictions. Comparing the point forecasts to their error-corrected probabilistic counterparts, it can be seen from the constant rate forecasts that its large mispredictions are less pronounced when using a rolling correction period. However, for the no-change forecast, the *ecp* no-change forecast worsens 9-month horizon forecasts.



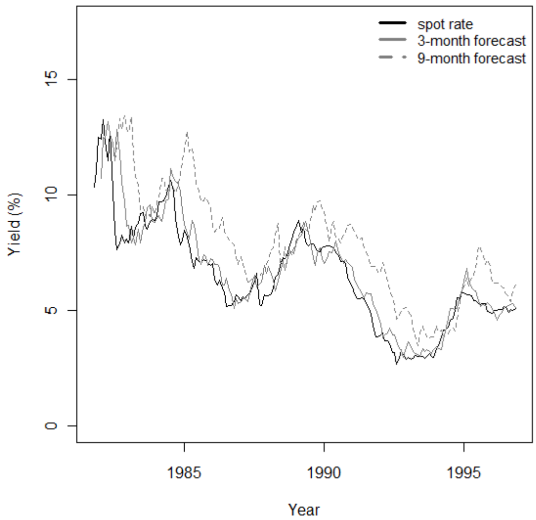
A.1.1 No-change forecast



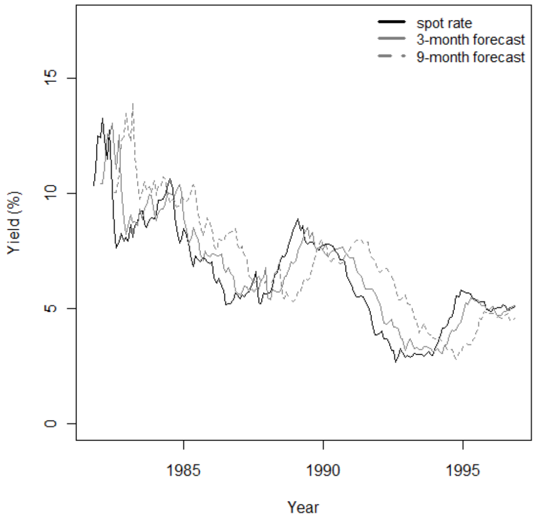
A.1.2 Constant rate forecast



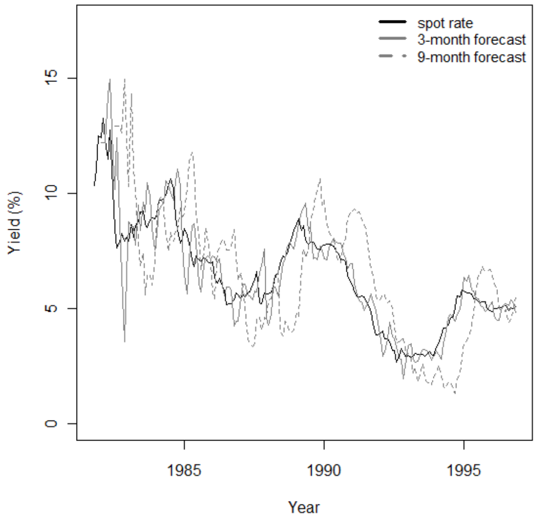
A.1.3 Autoregressive forecast



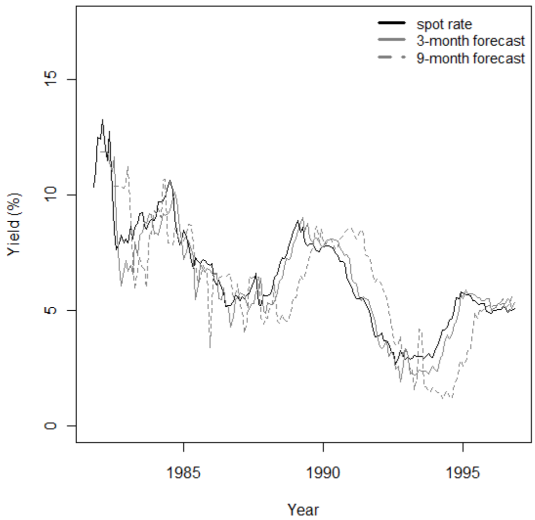
A.1.4 Forward rate forecast



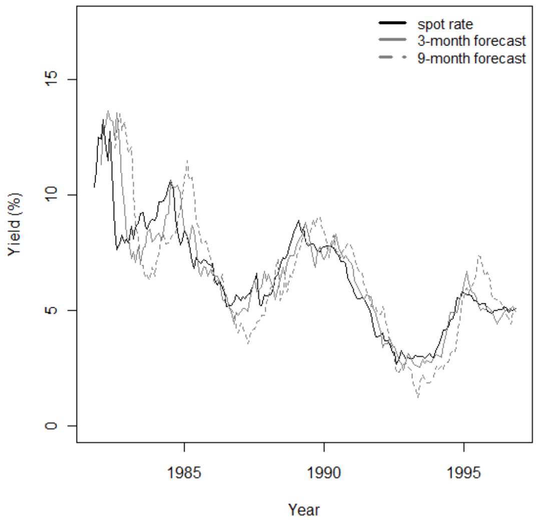
A.1.5 Ecp no-change forecast



A.1.6 Ecp constant rate forecast



A.1.7 Ecp autoregressive forecast



A.1.8 Ecp forward rate forecast

B. Evaluation separated by horizons

To gain further insight into the evaluation of the forecasts under each scoring function, the following figures display performances with each forecast horizon treated separately. The function $S(k)$ evaluates a forecast's prediction for horizon k. For the average directional accuracy, $S(k)$ takes the form

$$S_{ADA}(k) = \frac{1}{n_k} \sum_{t=0}^{n_k-1} \left(\mathbf{1}\{y_{t+k}^{t+1} \geq y_{t+k}^{t}\} \mathbf{1}\{\hat{x}_{t+k}^{t} \geq y_{t+k}^{t}\} + \mathbf{1}\{y_{t+k}^{t+1} < y_{t+k}^{t}\} \mathbf{1}\{\hat{x}_{t+k}^{t} < y_{t+k}^{t}\} \right).$$

Results are displayed in Figure B.1.1 to Figure B.1.8.

Similarly, the mean absolute error is evaluated for each forecast horizon by

$$S_{MAE}(k) = \frac{1}{n_k} \sum_{t=0}^{n_k-1} |\hat{x}_{t+k}^{t} - y_{t+k}^{t+1}|$$

and results are displayed in Figure B.2.1 to Figure B.2.8.

For the root-mean-squared error, $S(k)$ takes the form

$$S_{RMSE}(k) = \sqrt{\frac{1}{n_k} \sum_{t=0}^{n_k-1} (\hat{x}_{t+k}^{t} - y_{t+k}^{t+1})^2}$$

and results are displayed in Figure B.3.1 to Figure 3.1.8.

Profit rule A is evaluated with every forecast horizon treated separately by

$$\begin{aligned} S_{\pi_A}(k) = \frac{1}{m} \sum_{t=0}^{n_k} (&\mathbf{1}\{\hat{x}_{t+k}^{t} > x_t\}(x_{t+k} - y_{t+k}^{t}) \\ &+ \mathbf{1}\{\hat{x}_{t+k}^{t} < x_t\}(y_{t+k}^{t} - x_{t+k}) \\ &- \mathbf{1}\{\hat{x}_{t+k}^{t} \neq x_t\}(\$40)), \end{aligned}$$

where m is the duration of the forecasting period in years. Results are displayed from Figure B.4.1 to Figure B.4.8.

For profit rule B, $S(k)$ takes the form

$$\begin{aligned} S_{\pi_B}(k) = \frac{1}{m}\sum_{t=0}^{n_k} &(\mathbf{1}\{\hat{x}_{t+k}^t > y_{t+k}^t\}(x_{t+k} - y_{t+k}^t) \\ &+ \mathbf{1}\{\hat{x}_{t+k}^t < y_{t+k}^t\}(y_{t+k}^t - x_{t+k}) \\ &- \mathbf{1}\{\hat{x}_{t+k}^t \neq y_{t+k}^t\}(\$40)) \end{aligned}$$

and results can be seen in Figure B.5.1 through Figure B.5.8.

Profit rule C has $S(k)$ function

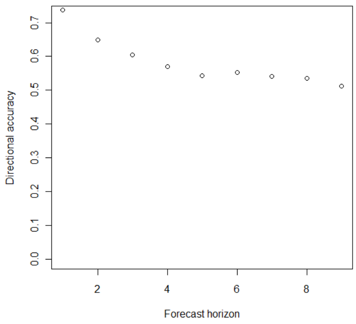
$$\begin{aligned} S_{\pi_C}(k) = \frac{1}{m}\sum_{t=0}^{n_k} \Big(&\mathbf{1}(\hat{x}_{t+k}^t \neq x_t)(\mathbf{1}\{\hat{x}_{t+k}^t > y_{t+k}^t\}(x_{t+k} - y_{t+k}^t) \\ &+ \mathbf{1}\{\hat{x}_{t+k}^t < y_{t+k}^t\}(y_{t+k}^t - x_{t+k}) \\ &- \mathbf{1}\{\hat{x}_{t+k}^t \neq y_{t+k}^t\}(\$40))\Big), \end{aligned}$$

results shown from Figure B.6.1 to Figure B.6.8.

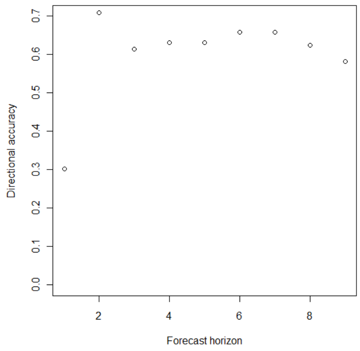
Profit rule D has $S(k)$ function

$$\begin{aligned} S_{\pi_D}(k) = \frac{1}{m}\sum_{t=0}^{n_k} &(\mathbf{1}\{\hat{x}_{t+k}^t > x_t\}\mathbf{1}\{y_{t+k}^t < x_t\}(x_{t+k} - y_{t+k}^t) \\ &+ \mathbf{1}\{\hat{x}_{t+k}^t < x_t\}\mathbf{1}\{y_{t+k}^t > x_t\}(y_{t+k}^t - x_{t+k}) \\ &- \mathbf{1}\{\hat{x}_{t+k}^t \neq y_{t+k}^t\}(\$40)) \end{aligned}$$

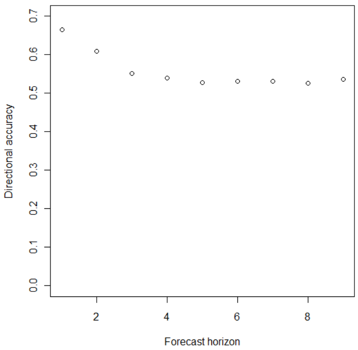
with results displayed from Figure B.7.1 to Figure B.7.8. The figures reveal the varying strengths of the forecasts over different forecast horizons. Notice the no-change forecast's strong performance over long forecast horizons. The erratic performance of the forward rate forecast seems obscure and is possibly due to a flawed modeling of the yield curve as discussed in section 6.



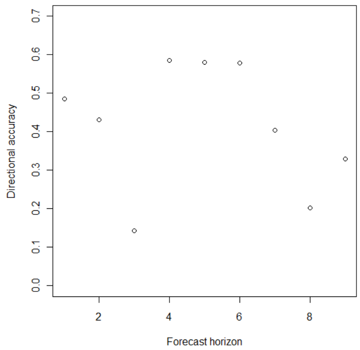
B.1.1 Naive no-change



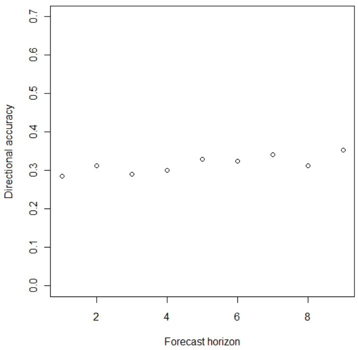
B.1.2 Constant rate



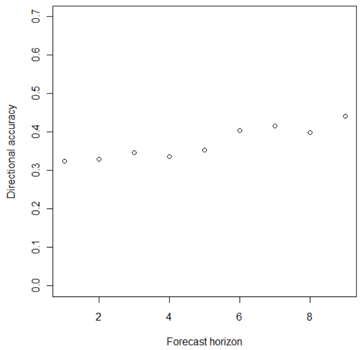
B.1.3 Autoregressive



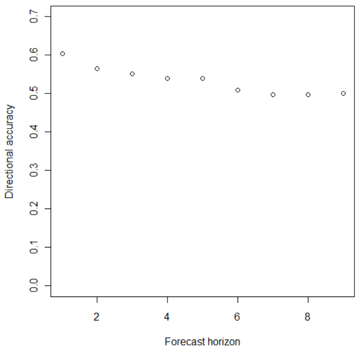
B.1.4 Forward rate



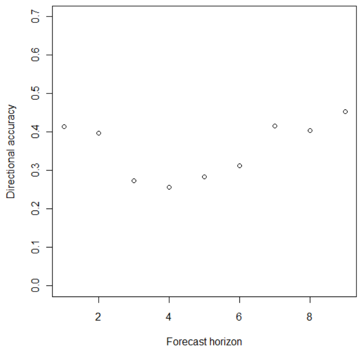
B.1.5 Ecp naive no-change



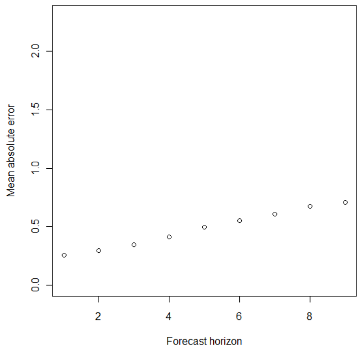
B.1.6 Ecp constant rate



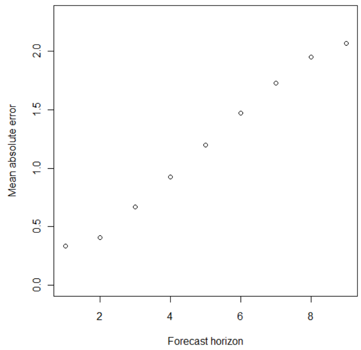
B.1.7 Ecp autoregressive



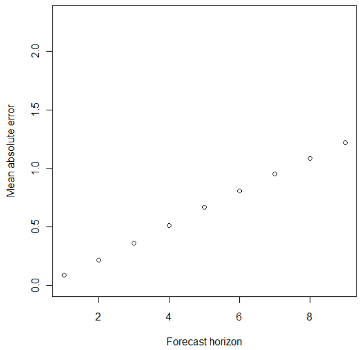
B.1.8 Ecp forward rate



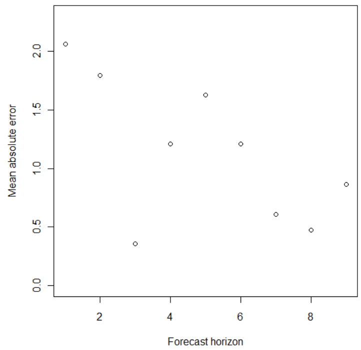
B.2.1 Naive no-change



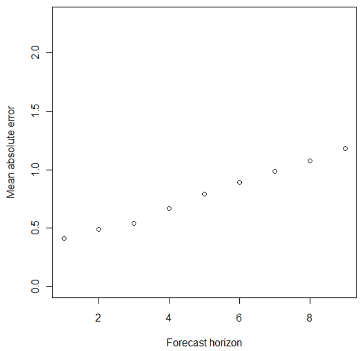
B.2.2 Constant rate



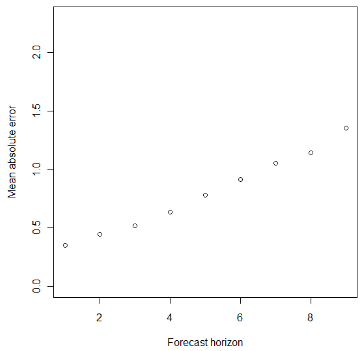
B.2.3 Autoregressive



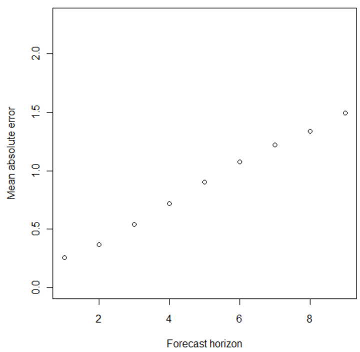
B.2.4 Forward rate



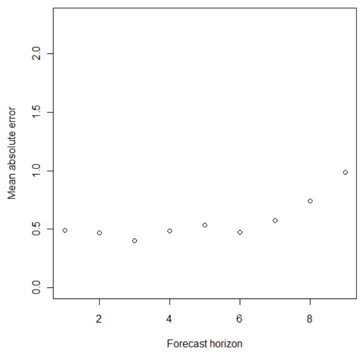
B.2.5 Ecp no-change forecast



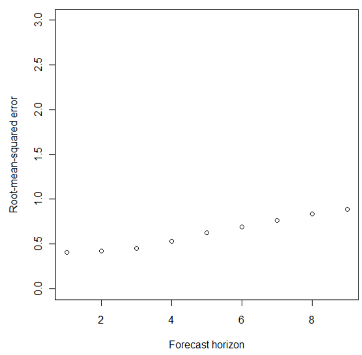
B.2.6 Ecp constant rate forecast



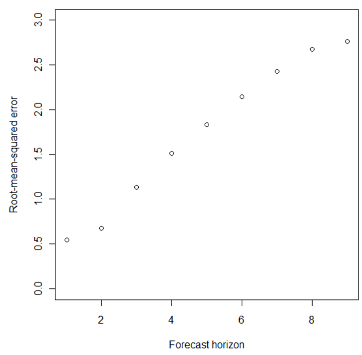
B.2.7 Ecp autoregressive



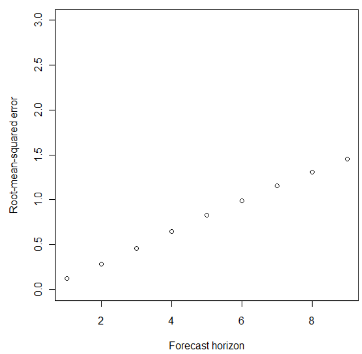
B.2.8 Ecp forward rate



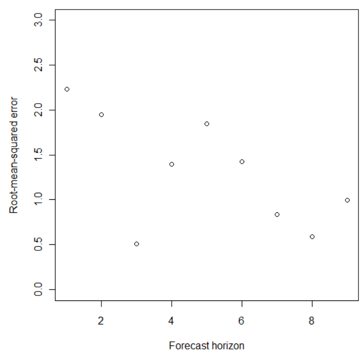
B.3.1 Naive no-change



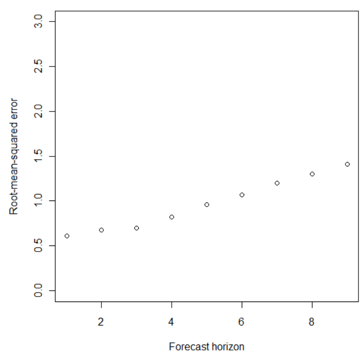
B.3.2 Constant rate



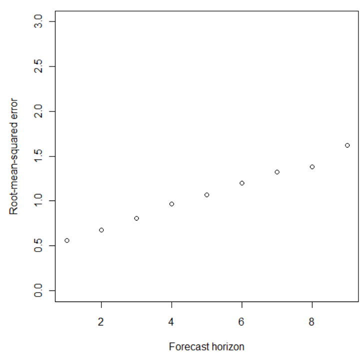
B.3.3 Autoregressive



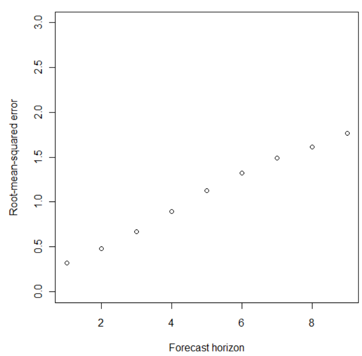
B.3.4 Forward rate



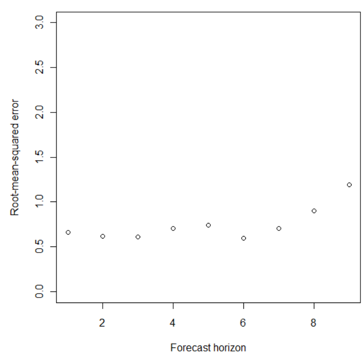
B.3.5 Ecp no-change forecast



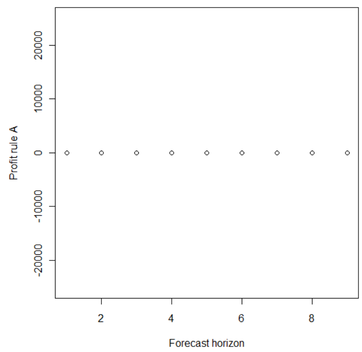
B.3.6 Ecp constant rate forecast



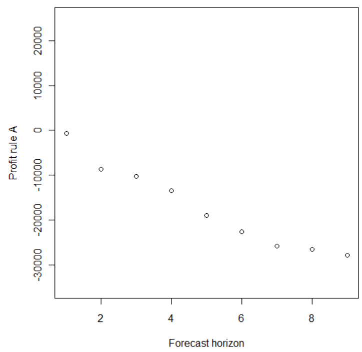
B.3.7 Ecp autoregressive



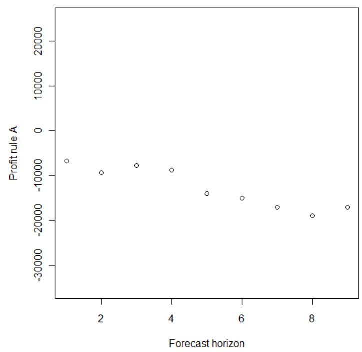
B.3.8 Ecp forward rate



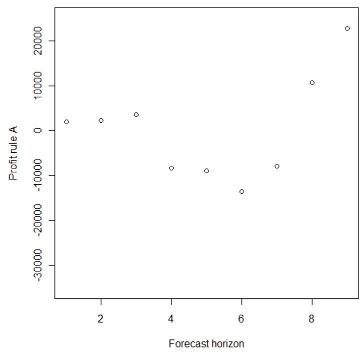
B.4.1 Naive no-change



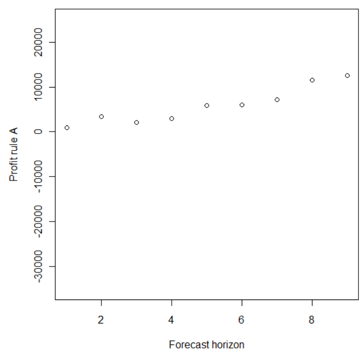
B.4.2 Constant rate



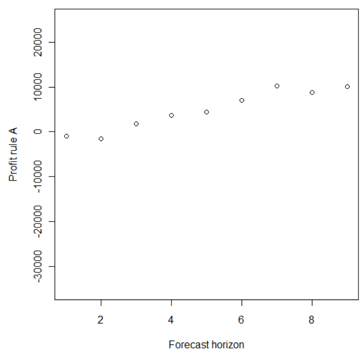
B.4.3 Autoregressive



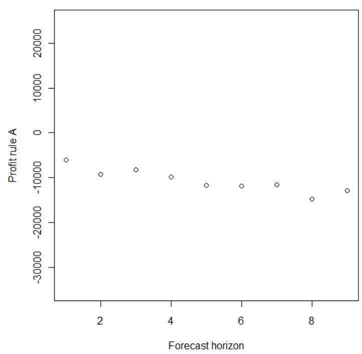
B.4.4 Forward rate



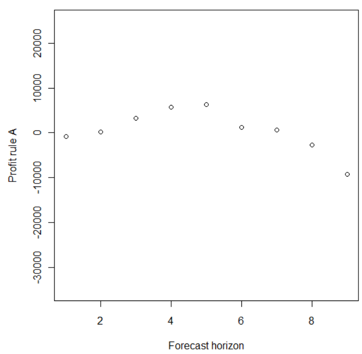
B.4.5 Ecp no-change forecast



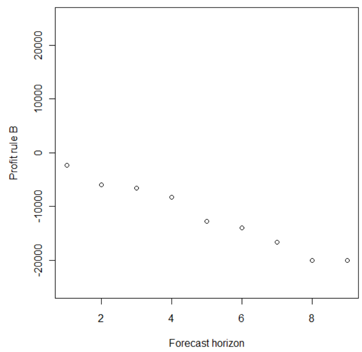
B.4.6 Ecp constant rate forecast



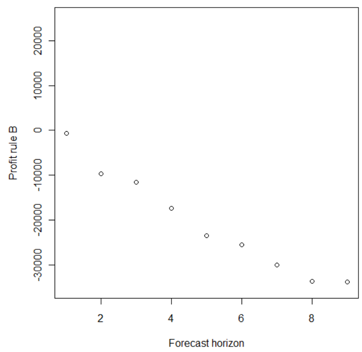
B.4.7 Ecp autoregressive



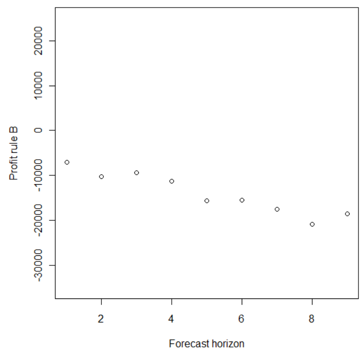
B.4.8 Ecp forward rate



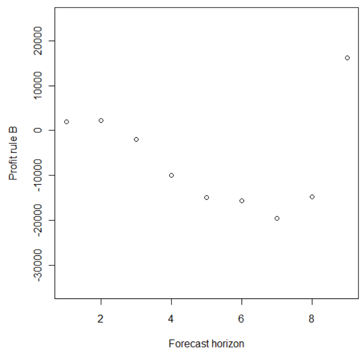
B.5.1 Naive no-change



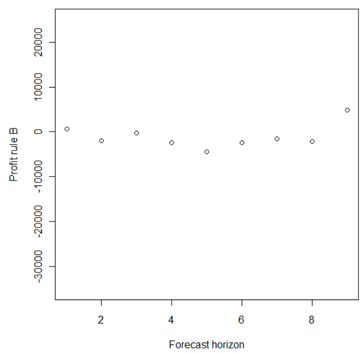
B.5.2 Constant rate



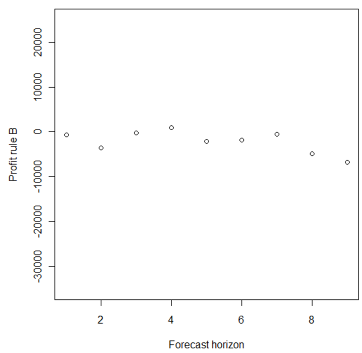
B.5.3 Autoregressive



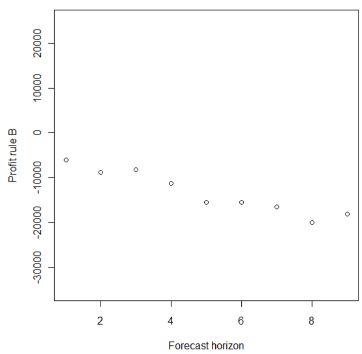
B.5.4 Forward rate



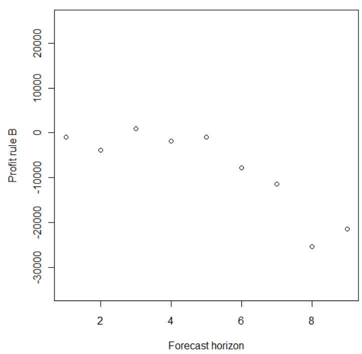
B.5.5 Ecp no-change forecast



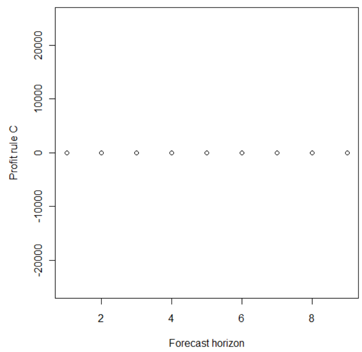
B.5.6 Ecp constant rate forecast



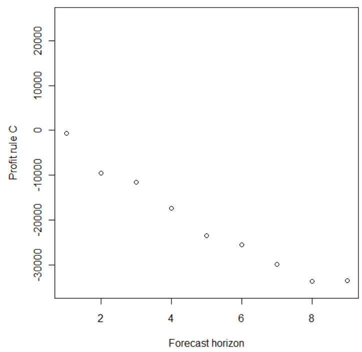
B.5.7 Ecp autoregressive



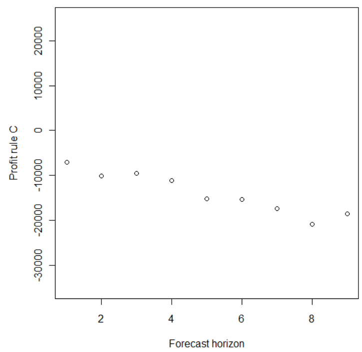
B.5.8 Ecp forward rate



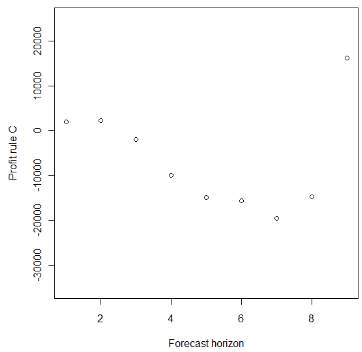
B.6.1 Naive no-change



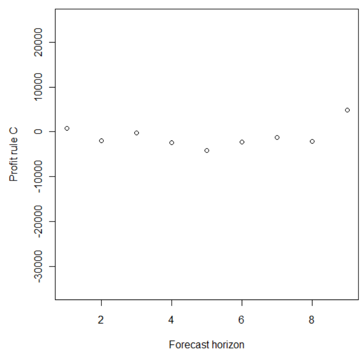
B.6.2 Constant rate



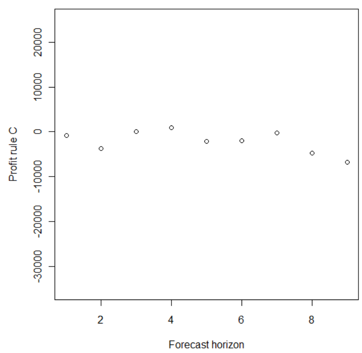
B.6.3 Autoregressive



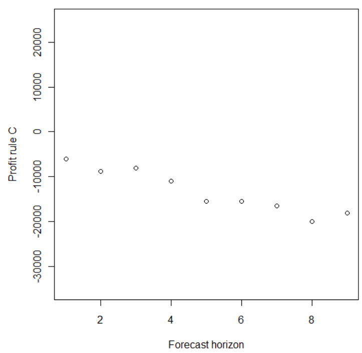
B.6.4 Forward rate



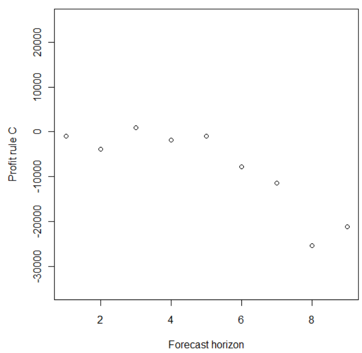
B.6.5 Ecp no-change forecast



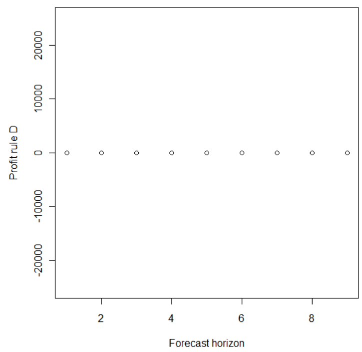
B.6.6 Ecp constant rate forecast



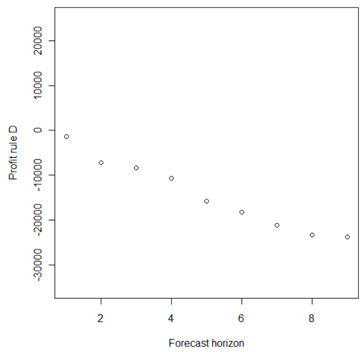
B.6.7 Ecp autoregressive



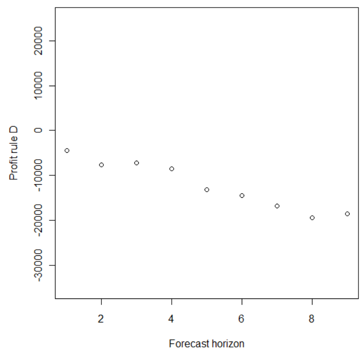
B.6.8 Ecp forward rate



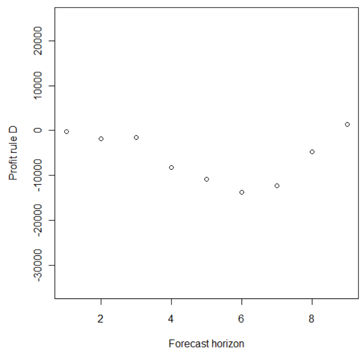
B.7.1 Naive no-change



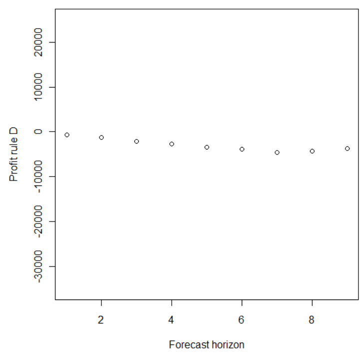
B.7.2 Constant rate



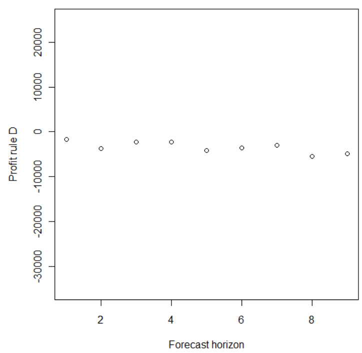
B.7.3 Autoregressive



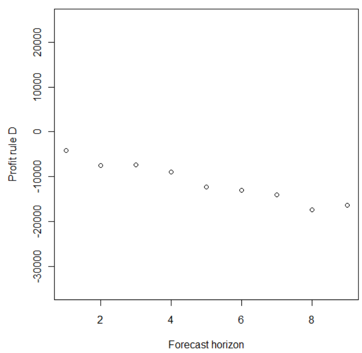
B.7.4 Forward rate



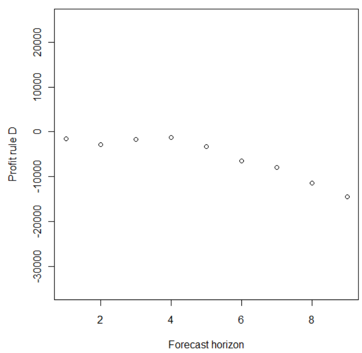
B.7.5 Ecp no-change forecast



B.7.6 Ecp constant rate forecast



B.7.7 Ecp autoregressive

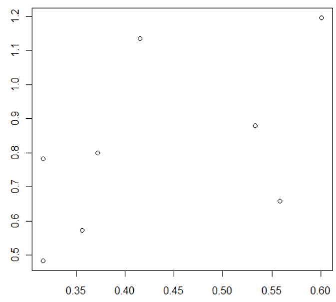


B.7.8 Ecp forward rate

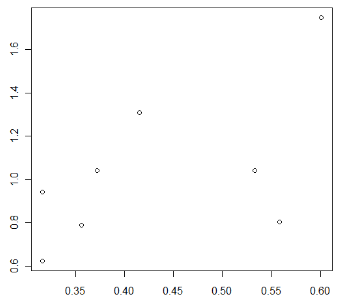
C. Scatter plots

C.1. Averaged by month and prediction horizon

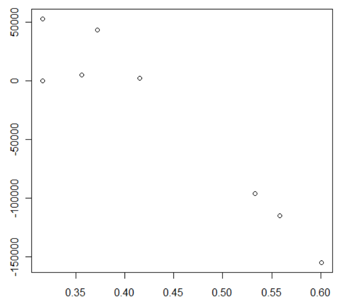
In this section, forecasts' predictions are plotted averaged over both the months and prediction horizons. Each point corresponds to one forecast. There are 8 points per plot, representing the 8 forecasts made. The first scoring function named in the caption is plotted on the horizontal axis, the second is plotted on the vertical axis.



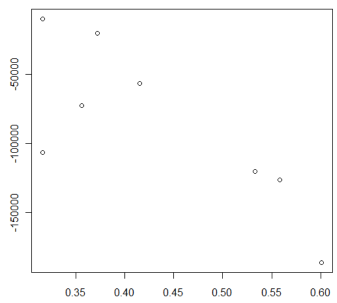
C.1.1 DA vs. MAE



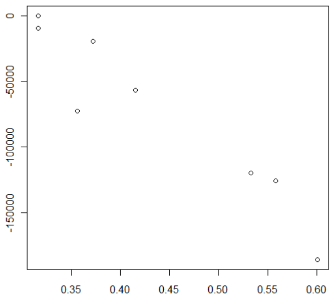
C.1.2 DA vs. RMSE



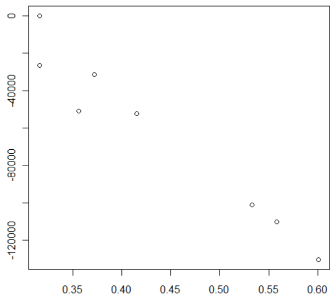
C.1.3 DA vs. Profit A



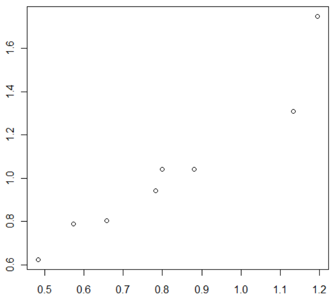
C.1.4 DA vs. Profit B



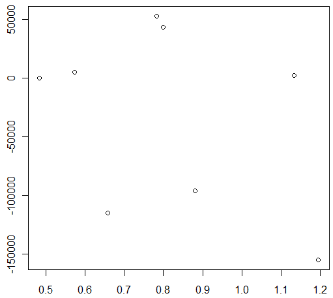
C.1.5 DA vs. Profit C



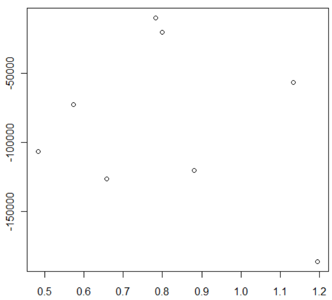
C.1.6 DA vs. Profit D



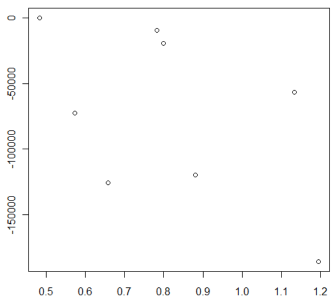
C.1.7 MAE vs. RMSE



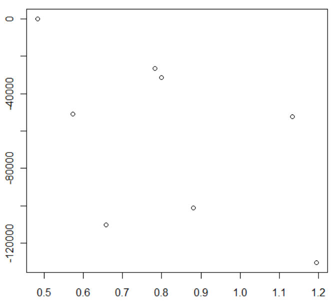
C.1.8 MAE vs. Profit A



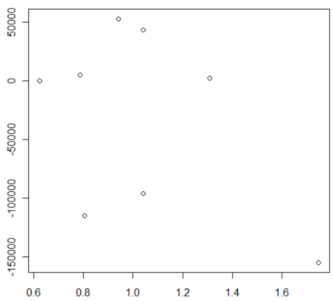
C.1.9 MAE vs. Profit B



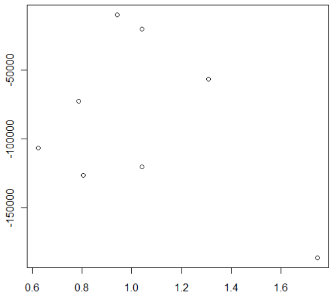
C.1.10 MAE vs. Profit C



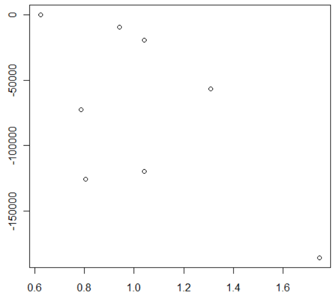
C.1.11 MAE vs. Profit D



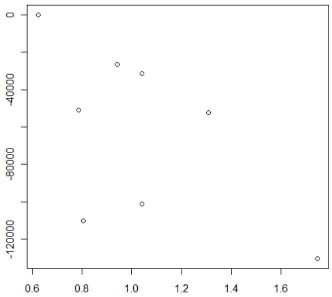
C.1.12 RMSE vs. Profit A



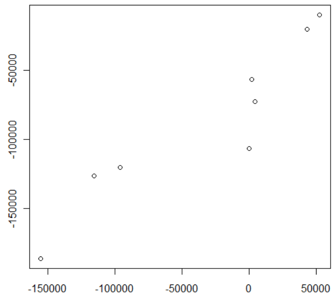
C.1.13 RMSE vs. Profit B



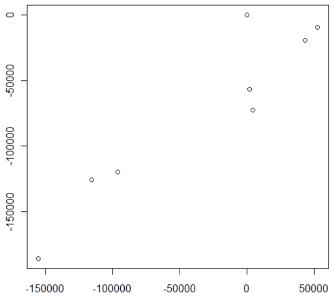
C.1.14 RMSE vs. Profit C



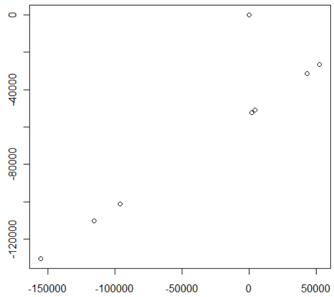
C.1.15 RMSE vs. Profit D



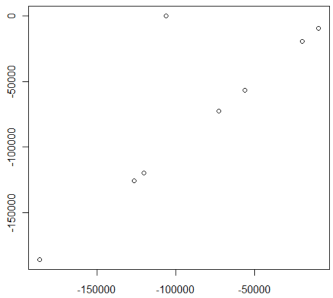
C.1.16 Profit A vs. Profit B



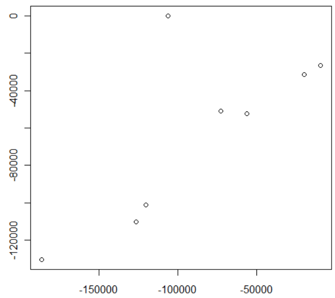
C.1.17 Profit A vs. Profit C



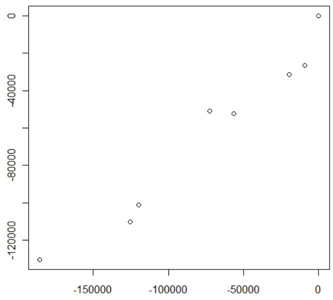
C.1.18 Profit A vs. Profit D



C.1.19 Profit B vs. Profit C



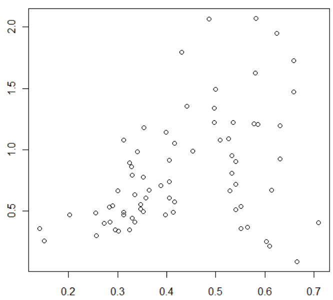
C.1.20 Profit B vs. Profit D



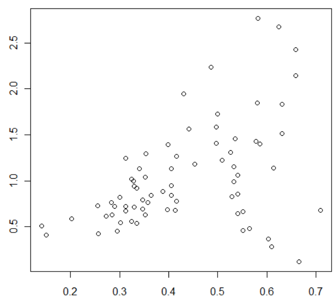
C.1.21 Profit C vs. Profit D

C.2. Averaged by month

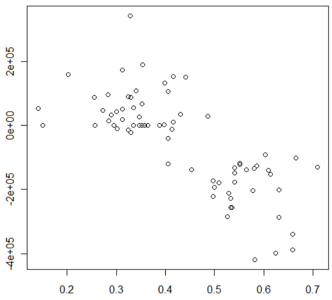
In this section, the performance of a forecast is averaged over every month while the forecast horizons are treated separately. This method yields 9 performance measurements per forecast, thus totaling 72 points per plot. The first scoring function named in the caption is plotted on the horizontal axis, the second is plotted on the vertical axis.



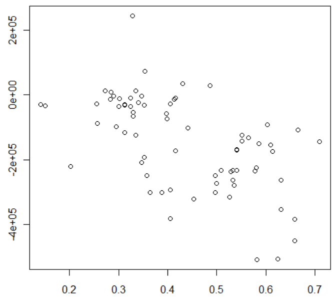
C.2.1 DA vs. MAE



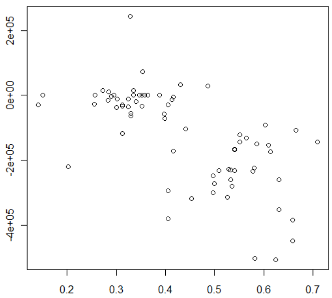
C.2.2 DA vs. RMSE



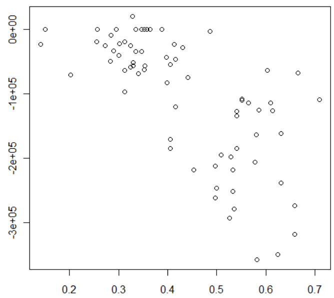
C.2.3 DA vs. Profit A



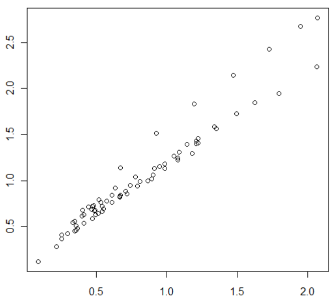
C.2.4 DA vs. Profit B



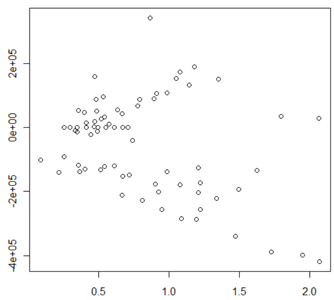
C.2.5 DA vs. Profit C



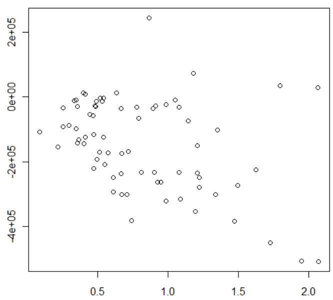
C.2.6 DA vs. Profit D



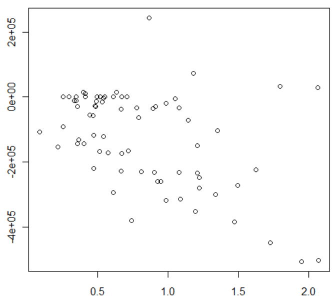
C.2.7 MAE vs. RMSE



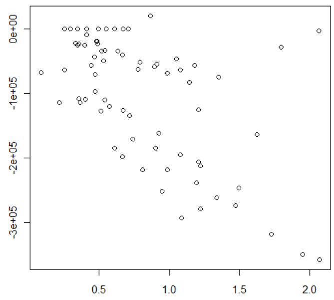
C.2.8 MAE vs. Profit A



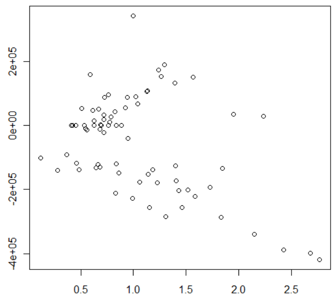
C.2.9 MAE vs. Profit B



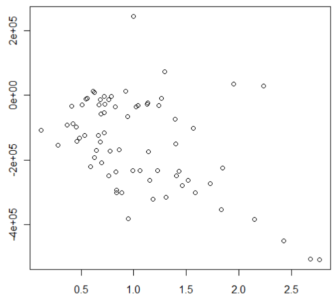
C.2.10 MAE vs. Profit C



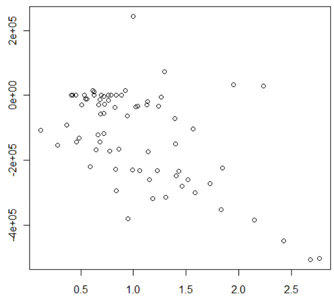
C.2.11 MAE vs. Profit D



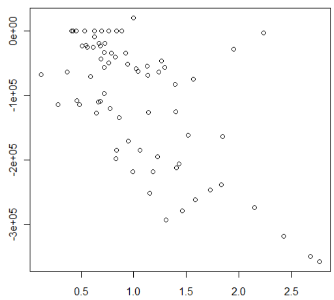
C.2.12 RMSE vs. Profit A



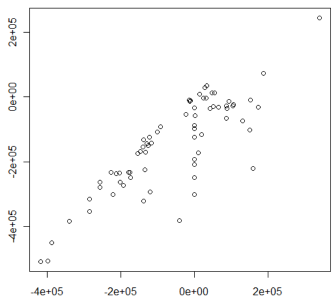
C.2.13 RMSE vs. Profit B



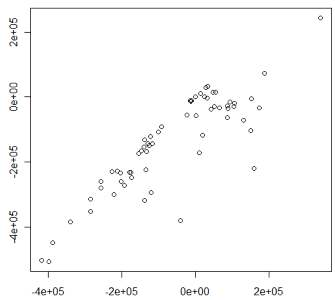
C.2.14 RMSE vs. Profit C



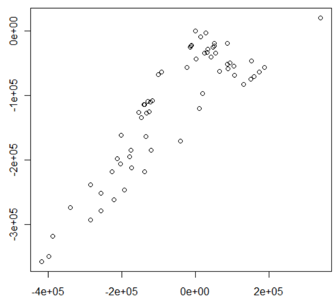
C.2.15 RMSE vs. Profit D



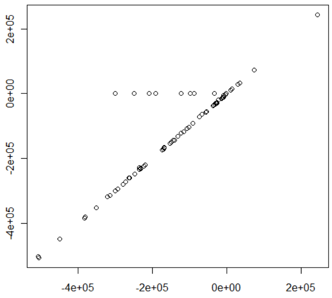
C.2.16 Profit A vs. Profit B



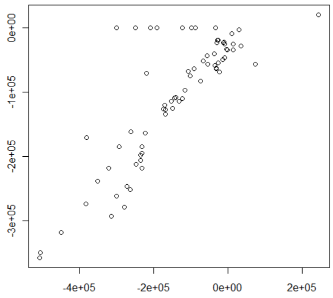
C.2.17 Profit A vs. Profit C



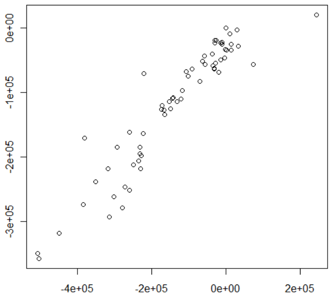
C.2.18 Profit A vs. Profit D



C.2.19 Profit B vs. Profit C



C.2.20 Profit B vs. Profit D



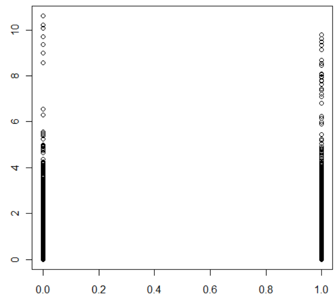
C.2.21 Profit C vs. Profit D

C.3. Individual predictions

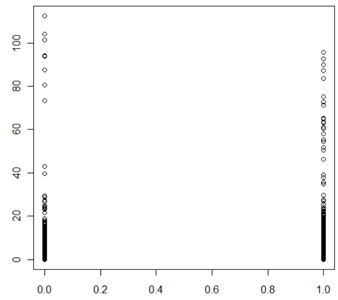
In this section, each individual prediction's performance is plotted. Every forecast makes 1581 predictions, equaling 12648 points per plot. The first scoring function named in the caption is plotted on the horizontal axis, the second is plotted on the vertical axis. Since individual predictions are plotted, summary measures are replaced by their respective scoring functions, changing the description from the MAE and RMSE to absolute error (AE) and squared error (SE) respectively.

Note from Figure C.3.7 that not all points lie on a deterministic line. This is due to the probabilistic forecasts delivering different point predictions under the absolute error and the squared error scoring functions.

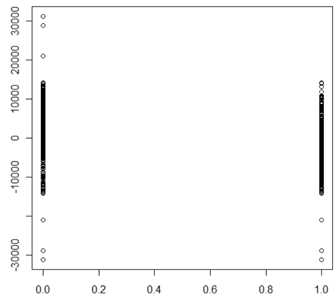
The general "cross" shape of Figures C.3.16 to C.3.21 can be explained in that the profit rules only differ in whether a futures contract is bought or sold. The magnitude of the resulting profit (or loss) will be the same, hence all points lie on the diagonals of the plane. Exceptions are when no action is taken under one of the profit rules in which case the points will lie on the horizontal or vertical axis.



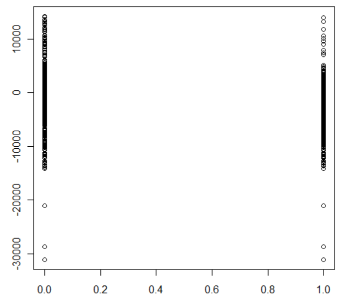
C.3.1 DA vs. AE



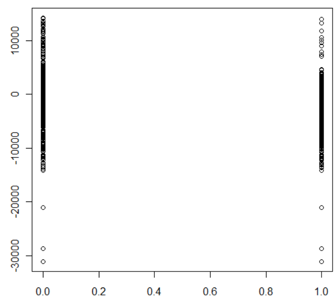
C.3.2 DA vs. SE



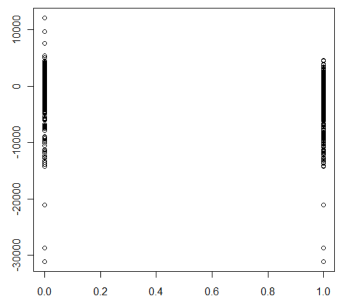
C.3.3 DA vs. Profit A



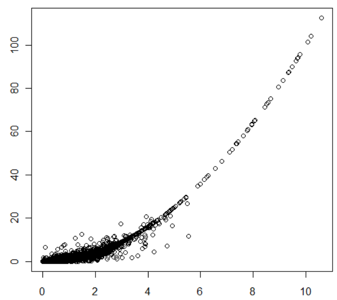
C.3.4 DA vs. Profit B



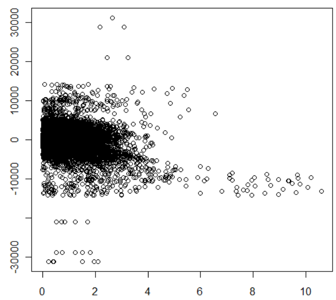
C.3.5 DA vs. Profit C



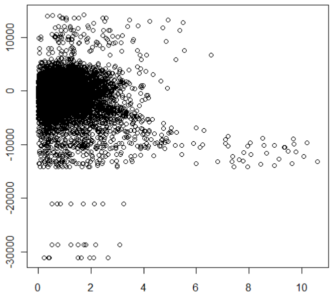
C.3.6 DA vs. Profit D



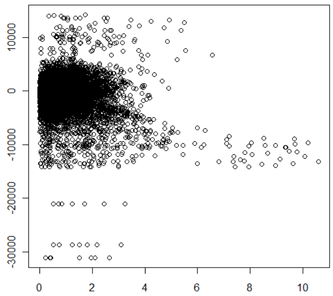
C.3.7 AE vs. SE



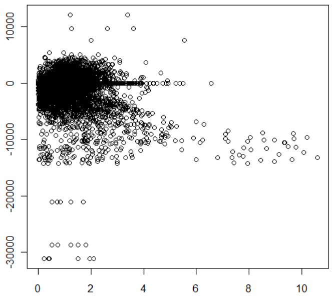
C.3.8 AE vs. Profit A



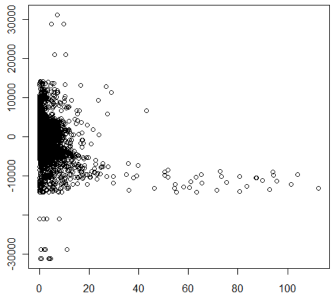
C.3.9 AE vs. Profit B



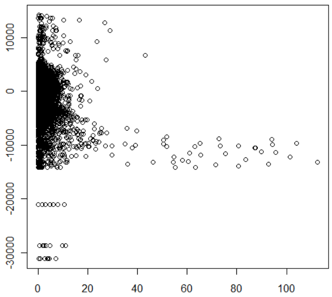
C.3.10 AE vs. Profit C



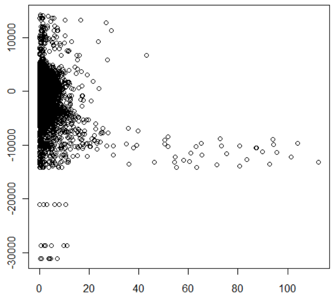
C.3.11 AE vs. Profit D



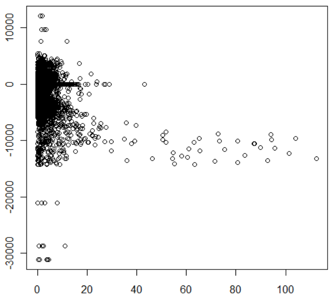
C.3.12 SE vs. Profit A



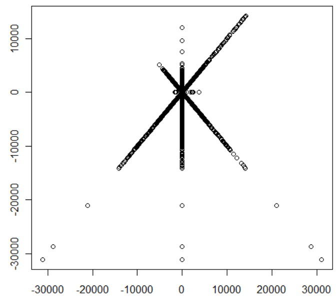
C.3.13 SE vs. Profit B



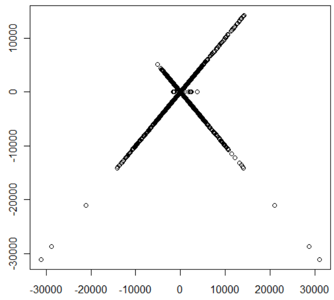
C.3.14 SE vs. Profit C



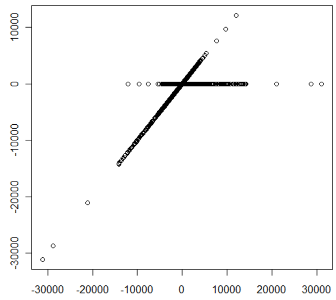
C.3.15 SE vs. Profit D



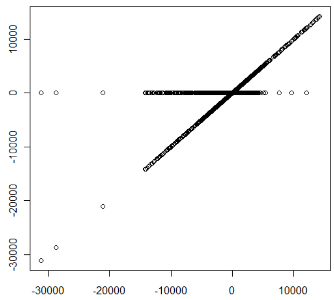
C.3.16 Profit A vs. Profit B



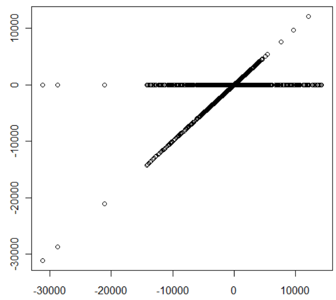
C.3.17 Profit A vs. Profit C



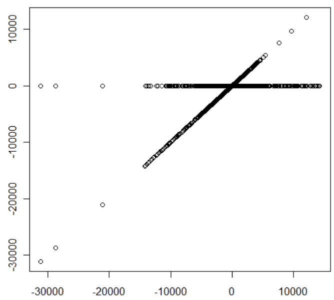
C.3.18 Profit A vs. Profit D



C.3.19 Profit B vs. Profit C



C.3.20 Profit B vs. Profit D



C.3.21 Profit C vs. Profit D

D. R Code

D.1. AR fitting and predicting

What follows is the R code used for fitting an autoregressive model to the data and making predictions, as explained in section 4.1.3. In particular, it relies on the implementation of the "tseries" R package, see Trapletti & Hornik (2011)

```
\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\# \\

library(tseries) \\

\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\# \\

ar.prediction <- function(x,n) \{ \\
\# ar.prediction predicts the next 9 forecast horizons based on a fitted
autoregressive model. The function performs a stationarity test on a rolling
training period of length n and fits either a AR(1) or ARIMA(1,1,0) model \\
\# x is the data input, minimum length has to be n \\
\# n is the desired length of the training period \\

if( length(x) < n) return(''the length of your input is too short!'') \\

data.length <- length(x) \\
stationarity.vector <- -999 \\

\# stationarity.vector is a vector that is filled with the entries
''stationary'' or ''non-stationary'' by the following function
stationarity.test; each entry corresponds to a training period. see
stationarity.test description for more info \\

stationarity.test <- function(x) \{ \\
if(kpss.test(x, null = ''Trend'')\$statistic > 0.146) stationarity.vector
<- ''non-stationary'' \\
else if (kpss.test(x, null = ''Trend'')\$statistic <= 0.146)
```

```
stationarity.vector <- ''stationary'' \\
return(stationarity.vector)\} \\

\# stationarity.test performs a kpss (Kwiatkowski-Phillips-Schmidt-Shin
tests) test for each training period to determine whether it is stationary or
non-stationary with a 95\% level of confidence. The results are saved in
stationarity.vector \\

for(i in 1:(data.length-n+1)) stationarity.vector[i]
<- stationarity.test(x[i:(i+n-1)]) \\

prediction <- -999 \\
prediction <- matrix(prediction, nrow = data.length-n+1, ncol = 9) \\

for(i in 1:(data.length-n+1))\{ \\
if (stationarity.vector[i] == ''stationary'') prediction[i,]
<- predict(ar(x[i:(i+n-1)], order.max = 1), n.ahead = 9)\$pred \\
else if (stationarity.vector[i] == ''non-stationary'') prediction[i,]
<- predict(arima(x[i:(i+n-1)], order = c(1,1,0), xreg=1:n), n.ahead = 9,
newxreg = (n+1):(n+9))\$pred \\
\}

\# The preceeding for-loop creates predictions from each training period's
fitted model to the following 9 forecast horizons. The results are saved in
the matrix ''prediction'' \\

prediction <- round(prediction, digits = 2) \\
prediction <- as.data.frame(prediction) \\

\# Note: the last row of the matrix ''prediction'' can generally be discarded
since it represents a prediction made at the last point in time where
predictions are no longer necessary \\

return(prediction)\} \\

\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\# \\

realization <- function(x)\{ \\
\# This function turns a vector of data into a matrix that can be used as a
realization matrix with which forecasts are compared \\

data.length <- length(x) \\

realization <- NA \\
```

```
realization <- matrix(realization, nrow = data.length, ncol = 9) \\

for(i in 1:(data.length-8)) realization[i,] <- x[i:(i+8)] \\
for(j in (data.length-7):data.length) realization[j,1:(data.length-j+1)]
<- x[j:data.length] \\


realization <- as.data.frame(realization) \\

return(realization)\} \\

\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\# \\

test.n <- function(x,m)\{ \\
\# Tests for which training period the mean average error is smallest \\
\# x is the data to perform the predictions on \\
\# m is a vector of different training periods to compare
(e.g. m = c(12,18,24,36) ) \\

data.length <- length(x) \\
r <- realization(x) \\
a <- -999 \\

\# Vector 'a' will store the results of each fitting's mean average error \\

for(n in m)\{ \\
p <- ar.prediction(x,n) \\

if(n==max(m))\{ a[which(m==n)] <- sum(abs(p - r[-(1:(max(m)-1)),]),
na.rm = TRUE)/((data.length - max(m) + 1 - 8)*9 + 36) \} \\
else\{ a[which(m==n)] <- sum(abs(p[-(1:(max(m)-n)),] - r[-(1:(max(m)-1)),]),
na.rm = TRUE)/((data.length - max(m) + 1 - 8)*9 + 36) \} \\

\} \\

\# One has to be careful that over the same set of data, model fittings with
longer training periods produce fewer forecasts than AR fittings with shorter
training periods; however, when comparing forecasts of different fittings, all
should be evaluated only over the realizations which are common to all (i.e.
the realizations of the fitting with the longest training period) \\

return(a)\}
```

D.2. Ecp forecast

The following code is used for turning a forecast into an error-corrected probabilistic forecast as explained in section 4.2:

```
\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\# \\

pointforecast <- function(x, fct)\{ \\
\# This function derives a 'best' point forecast based on a probability
forecast \\
\# x is the probability forecast matrix \\
\# fct is the function to apply to retrieve a point forecast
(e.g. mean, median) \\

m <- dim(x)[2]/9 \\
point <- numeric() \\

for(i in 1:9)\{ \\
a <- apply(x[,(1:m) + (i-1)*m], 1, fct) \\
point <- round(cbind(point, a), digits = 2) \\
\} \\

return(point)\} \\

\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\# \\

ecpforecast <- function(data, pf, m)\{ \\
\# Creates a matrix that serves as the probability forecast \\
\# See ecpf description for more info \\

a <- rbind( ecpf(data,pf,m,1) , matrix(NA, ncol = m, nrow = 0) ) \\
b <- rbind( ecpf(data,pf,m,2) , matrix(NA, ncol = m, nrow = 1) ) \\
c <- rbind( ecpf(data,pf,m,3) , matrix(NA, ncol = m, nrow = 2) ) \\
d <- rbind( ecpf(data,pf,m,4) , matrix(NA, ncol = m, nrow = 3) ) \\
e <- rbind( ecpf(data,pf,m,5) , matrix(NA, ncol = m, nrow = 4) ) \\
f <- rbind( ecpf(data,pf,m,6) , matrix(NA, ncol = m, nrow = 5) ) \\
g <- rbind( ecpf(data,pf,m,7) , matrix(NA, ncol = m, nrow = 6) ) \\
h <- rbind( ecpf(data,pf,m,8) , matrix(NA, ncol = m, nrow = 7) ) \\
i <- rbind( ecpf(data,pf,m,9) , matrix(NA, ncol = m, nrow = 8) ) \\

x <- cbind(a,b,c,d,e,f,g,h,i) \\
return(x)\} \\
```

```
\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\# \\

ecpf <- function(data,pf,m,k)\{ \\
\# Function creates an error-corrected probability forecast matrix based
on a point

forecasts matrix \\
\# data denotes the data for the training period as well as the data that
is to be

predicted (variable should be a vector) \\
\# pf denotes the point forecast matrix for the (m+1)th data entry to the
last data

entry (i.e. not including the initial correction period) \\
\# m denotes the length of the correction period \\
\# k denotes the forecast horizon \\

\# Note: the ''data'' input has to be modified to fit the point forecast
by which it is

supposed to be predicted: discard the beginning entries of the ''data''
vectors until

its length is +1 the number of rows of the point forecast matrix (see
below for more

info) \\

real <- as.matrix(realization(data)) \\
error <- pf - real[-1,] \\

\# The error of a point forecast. The first row is discarded from the
realization

matrix because it corresponds to the first ''input for the predictions''
which cannot be

part of their prediction \\

prob <- numeric() \\
```

```
for(i in 1:(dim(pf)[1]-m-k+1))\{ \\

e <- error[(i:(i+m-1)),k] \\
prob[(1:m) + (i-1)*m] <- sort( pf[m+i,k] + e) \\
\} \\

prob <- matrix(prob, ncol = m, byrow = TRUE) \\
return(prob)\} \\

\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\# \\

realization <- function(x)\{ \\
\# This function turns a vector of data into a matrix that can be used
as a realization \\

matrix with which forecasts are compared \\

data.length <- length(x) \\

realization <- NA \\
realization <- matrix(realization, nrow = data.length, ncol = 9) \\

for(i in 1:(data.length-8)) realization[i,] <- x[i:(i+8)] \\
for(j in (data.length-7):data.length) realization[j,1:(data.length-j+1)]
<- x \\

[j:data.length] \\


realization <- as.data.frame(realization) \\

return(realization)\} \\

\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\#\# \\

naive <- function(data)\{ \\
\# creates a naive no-change forecast matrix that can quickly be used
as a sample forecast\\

datam <- matrix(data[-length(data)], nrow = length(data) - 1, ncol = 9) \\
for(j in (length(data)-8):(length(data)-1)) datam[j,((length(data)-j+1):9)]
<- NA \\
return(datam)\} \\
```

Hiermit versichere ich, dass ich meine Arbeit selbstständig unter Anleitung verfasst habe, dass ich keine anderen als die angegebenen Quellen als Hilfsmittel benutzt habe, und dass ich alle Stellen, die dem Wortlaut oder dem Sinne nach anderen Werken entlehnt sind, durch die Angabe der Quellen als Entlehnung kenntlich gemacht habe.