Forward Rates:
Predictive Power and Trading Strategies

Schriftliche Promotionsleistung

zur Erlangung des akademischen Grades

Doctor rerum politicarum

vorgelegt und angenommen

an der Fakultät für Wirtschaftswissenschaft

der Otto-von-Guericke-Universität Magdeburg

Verfasser:	Diana Afanasenko
Geburtsdatum und -ort:	26. August 1982 in Nowosibirsk
Arbeit eingereicht am:	18. Oktober 2012

Gutachter der schriftlichen Promotionsleistung:

Prof. Dr. Peter Reichling

Prof. Dr. Horst Gischer

Datum der Disputation: 17. Juni 2014

Contents

List of Tables

List of Figures

List of Abbreviations

ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
AMA	Advanced Measurement Approach
AR	Autoregression
ASF	Available Stable Funding
BH	Buy-and-Hold
BIA	Basic Indicator Approach
BIS	Bank for International Settlements
bps	Basis points
cdf	Cumulative Distribution Function
CI	Cointegration
CIV	Cointegrating Vector
CSR	Conditional Sharpe Ratio
CVaR	Conditional Value at Risk
DAX	German Stock Index
DF	Dickey-Fuller
DGP	Data Generating Process
ECM	Error Correction Model
EH	Expectations Hypothesis
EU	European Union
ERVaR	Excess Return on Value at Risk
Fed	Federal Reserve
FFSA	Federal Financial Supervisory Authority
FPE	Final Prediction Error
FSB	Financial Stability Board
FV	Face Value
GLS	Generalized Least Squares
HQ	Hannan-Quinn
iid	Independent and identically distributed
IRBA	Internal Ratings-Based Approach

IRR	Internal Rate of Return
JB	Jarque-Berra
KPSS	Kwiatkowski, Philips, Schmidt and Shean
LCR	Liquidity Coverage Ratio
LPM	Lower Partial Moment
LR	Likelihood Ratio
MAE	Mean Absolute Error
MaRisk	Minimum Requirements for Risk Management
MSR	Modified Sharpe Ratio
MVaR	Modified Value at Risk
NSFR	Net Stable Funding Ratio
NW	Newey-West
OLS	Ordinary Least Squares
PP	Philipp-Perron
RDYC	Rolling Down the Yield Curve
RMSE	Root Mean Squared Error
RYC	Riding the Yield Curve
SA	Standardized Approach
SD	Standard Deviation
SIC	Swartz Information Criterion
SIFI	Systemically Important financial Institution
UPM	Upper Partial Moment
UPR	Upside Potential Ratio
VAR	Vector Autoregression
VaR	Value at Risk
YTM	Yield to Maturity

List of Symbols

Δ	First-difference operator
α	Intercept term in a regression equation
$\tilde{\alpha}$	Cointegrating vector
β_i	Parameter in the forward rate models
β	Slope coefficient in the pure Expectations Hypothesis
$\tilde{\beta}$	Matrix of the adjustment coefficients
δ_0	Intercept term in the Error Correction Model
δ_1	Adjustment coefficient in the Error Correction Model
ε	Error term
φ_1	Parameter in the Nelson-Siegel model
η	Error term
$\lambda_{\max}$	Maximum eigenvalue statistics in Johansen cointegration test
λ_{trace}	Trace statistics in Johansen cointegration test
μ	Mean
σ	Standard deviation
τ	Target rate of return
ς	Decay parameter in the Nelson-Siegel model
ϑ	The slope coefficient in the Dickey-Fuller tests
AR(p)	Autoregression model of order p
b	Order of integration of a linear combination of two time series
C	Coupon payment
CI (d,b)	Cointegrated of order d,b
Cov	Covariance
d	Order of integration of a time series
E	Expectations operator
exr	Excess return
F	Value of F-statistics
$F(x)$	Cumulative probability function
f_i	Forward rate lagged by i periods

$f_t^{t+i;t+i+m}$	Forward interest rate determined at time *t* for the contact that starts in *t*+*i* and lasts for *m* periods
$f(x)$	Density function
$G(x)$	Cumulative probability function
g	Level of statistical significance
I(d)	Integrated of order *d*
I	Identity matrix
i	Holding period in the RYC strategy
JB	Jarque-Berra statistics
K	Kurtosis
k	Order of the VAR model
H	The number of regression coefficients in the Chow test
L^m	Liquidity premium for the maturity *m*
L_t^m	Liquidity premium for the maturity *m* determined at time *t*
L	Number of lags in VAR and Cointegration equation
LPM_q	Lower partial moment of degree *q*
m	Maturity
max	Maximum
N	Number of observations
n	Holding period
P	Price of a bond
P_t^m	Price of the bond with maturity *m* determined at time *t*
p-value	Probability value
p	The number of series in the CI model
q	Degree of the lower partial moment
R_i	Rate of return on the i^{th} investment
r	Number of cointegrating vectors
$r_{\mathrm{BE},t}$	Breakeven rate at time *t*
r^{RYC}	Rate of return on the RYC strategy
r^{RDYC}	Excess return on the RDYC strategy
r_t	Spot interest rate determined in period *t*
$\hat{r}_t$	Forecasted value of a spot interest rate determined in period *t*

r_f	Risk-free rate of return
r_{t+i}^{m}	Spot interest rate with maturity m determined in period $t+i$
S	Skewness
S_i	Sharpe Ratio
sr_m	Swap rate of an m-year swap
T	Sample size
T_1	Size of subsample 1
T_2	Size of subsample 2
t	Value of t-statistics
t	Time trend
U	Theil's inequality coefficient
UPM_q	Upper partial moment of degree q
Var	Variance
y	Yield to maturity
z	Drift term in the Dickey-Fuller test

1. Introduction

The functional relationship between the interest rates and the corresponding maturities is referred to as the term structure of interest rates. If observed for a significantly long period of time, this relationship possesses specific features. One of them is that, on average, long-term interest rates exceed the short-term interest rates and the shape of the term structure is upward-sloping. Although such a type of the term structure prevails, sometimes the shape of the term structure becomes flat and even downward-sloping. Empirical observations associated with the term structure gave rise to several term structure theories. While the first papers appeared in the 1940s of the 20th century, studies seeking to find evidence helping to explain the term structure can hardly be counted. As a result of this research, the expectations theory of the term structure emerged, containing three versions: the pure expectations theory, the liquidity theory, and the preferred habitat theory.

From the above three versions, the pure expectations hypothesis (EH) received by far the greatest attention in the academic literature. The pure expectations theory states that long-term interest rates are determined as an average of current and future expected short-term interest rates. Because it explains long-term interest rates relying exclusively on expectations, this version is often called an unbiased expectations theory. In contrast, the two other versions involve a risk-premium and are known as biased expectations theories. Due to its focus on expectations, the pure EH can explain every possible shape of the term structure, relating it to future expectations towards the future short-term interest rates. An important implication of the pure expectations theory is that forward rates are considered to be unbiased predictors of the future interest rates. During the past several decades, the pure expectations hypothesis has been subject to extensive research.

There are many reasons for the great interest in the term structure of interest rates. One of the most important aspects is its importance for monetary policy issues. Central banks can make use of different policy instruments to affect the short-term interest rate. However, investment and long-term consumption decisions are made based on the long-term interest rates. Thus, knowledge about how long- and short-term interest rates are related to each other would help to analyze the effectiveness of monetary policy issues. In the framework of the pure expectations theory, monetary authorities can only affect the long-term rates if they influence the expectations of market participants regarding the future short-term interest rates. Important implications from the observations of the term structure can also be made with respect to inflation expectations. An increase in the long-term interest rates is often interpreted as an indicator of a rise in expected inflation rates.

Factors that determine the term structure are also important for the government debt issuance. If long-term interest rates are an average of current and expected future short-term interest rates, it will be difficult to affect the term structure through buying and selling bonds of different maturities. In addition, the term structure is indispensible for pricing different financial instruments, such as bonds, swaps and interest rate options.

Studies of the term structure are especially important for forecasting purposes. If long-term interest rates represent an average of current and future expected short-term interest rates in accordance with the pure expectations theory, this result can be employed to predict the future short-term interest rates. Numerous empirical papers have focused on the information content of the term structure.

Despite a great number of empirical papers devoted to the pure expectations theory, an unambiguous conclusion still cannot be drawn. These studies use a variety of periods, interest rates, and apply different testing procedures. In general, the theory is rejected by the studies that used data for the US. For European data including Germany, the result is contradictory. Some authors strictly reject the theory whereas others find evidence in favor of the hypothesis. In general, evidence for Europe is more supportive than that for the US. The general failure of the theory gave rise to a discussion about the possible reasons for this rejection. They include measurement error, existence of a risk premium and an overreaction hypothesis.

An additional aspect to be mentioned is that testing the expectations theory in many of its forms is complicated by the fact that special assumptions regarding the expectations formation process are needed. In order to derive the testing equation, rational expectations are commonly assumed. Thus, the expectations theory may be only tested as a joint hypothesis, which results in ambiguity when interpreting test results. Either the expectations theory does not hold or the expectations are formed in a way different from the assumed one.

To overcome these difficulties, the empirical study presented in this dissertation thesis attempts to avoid the necessity of assuming some particular form of economic agents' expectations. Instead, the first research question of this study is whether forward rates of preceding periods contain any predictive power with respect to future spot rates. Within the numerous contributions on the term structure theories, tests applying UK or US data are clearly dominating. In this thesis, the existing literature on the German term structure will be extended employing recent data ranging from 1995 to 2007. In addition, due to recent developments in econometrics, new testing techniques have become available. In order to test the forecasting ability of lagged forward rates, cointegration analysis and the error correction model are employed.

The second research focus of this dissertation thesis is closely connected to and largely based on the results of the above analysis. If forward rates cannot serve as predictors of the future interest rates, some special implications can be drawn. For example, this could indicate that fixed income trading strategies that are based on the stable yield curve may be profitable. Among such strategies is the so called rolling down the yield curve (RDYC) strategy that involves borrowing short-term funds and investing them in long-term assets. This strategy represents a core business activity of banks, which roll over short-term funds in order to grant long-term loans. An alternative to this business activity is an investment on the capital market. The primary goal of this part of the analysis is to study and compare the performance of both strategies in Germany over the period from 1972 to 2007.

In Germany, RDYC is especially important to the savings bank group, representing a key business activity of this type of banks. In contrast, large commercial banks are especially active on the capital market, which represents the main source of income of this banking group. This becomes clear when considering the share of non-interest income and interest income of these banking groups. As focus on the client business has been under severe criticism because of concentration risks in the portfolio of such banks, an investigation of this issue can help to shed light on the success of two different business models.

In addition, the second strategy associated with a stable yield curve is the riding the yield curve (RYC) strategy. Although similar to the strategy above, it involves buying fixed income instruments and selling them prior to maturity. This strategy, although received some attention in the empirical literature, has been tested for Germany only once. In this thesis, this strategy will be considered and its performance compared with that of the buy-and-hold strategy. The purpose of this analysis is to draw further conclusions regarding the validity of the pure EH.

This dissertation thesis is structured in the following way. Chapter 2 aims at introducing the reader into the topic and giving the necessary prerequisites for the further analysis. In particular, section 2.1 addresses various methods that are commonly used in order to estimate the yield curve, such as estimation from zero-coupon bonds and swap rates as well as some of the theoretical models. The following section 2.2 acquaints the reader with different variations of the expectations hypothesis of the term structure and consists of three parts: section 2.2.1 contains the pure expectations hypothesis, section 2.2.2 deals with the liquidity preference hypothesis whereas section 2.2.3 discusses the preferred habitat theory. A further theory of the term structure, namely, the market segmentations theory, is introduced in section 2.3. In order to proceed with the model covered in this dissertation thesis it is essential to reflect the empirical literature in this area at first. Consequently, chapter 2 is completed with an extensive review of the previous empirical findings corresponding to the expectations theory of the term structure.

Chapter 3 covers the empirical results of the selected models. It starts with section 3.1, which introduces the selected econometric methods and is further divided into three sections. Firstly, section 3.1.1 provides insights into the selected models and formulates the main objectives of the following empirical analysis. However, before starting the estimation of the equations, it is indispensible to determine the time series properties of spot and forward rates. The next section 3.1.2 contains the results of the preliminary data analysis. Finally, based on the findings presented in subsection 3.1.2, subsection 3.1.3 describes in details the selected methods of econometric analysis, namely, the cointegration and the error-correction model.

After the necessary methodology has been described, section 3.2 is entirely devoted to the empirical findings on the explanatory power of forward rates with respect to the future spot rates. At first, section 3.2.1 contains the results of cointegration properties of spot and corresponding forward rates and presents the parameter estimates. Section 3.2.1 gives insights into a short-term dynamic of the considered models with the help of the error-correction model. The estimated parameters are then used to build predictions of the spot rate and test an out-of-sample performance of each model (subsection 3.2.3). Finally, subsection 3.2.4 summarizes the findings and outlines the direction of the research issues considered in chapter 4. The main findings of the analysis performed in chapter 3 indicate that forward rates do not possess significant predictive power with respect to the future spot interest rates. Thus, it could be possible to use this result and build strategies based on a stable yield curve.

The entire chapter 4 is devoted to the yield curve trading strategies and is comprised of six subsections. It starts with a categorization of fixed income strategies in section 4.1. The key attention is given to the description of the two strategies based on the upward sloping, stable yield curve: RDYC and RYC strategies. The specifics of these strategies as well as return derivations are presented in subsection 4.1.3. The riding the yield curve strategy received significant attention of the academic community. In order to give an overview on the empirical performance of this strategy, major findings of previous research are summarized in section 4.2.

The following sections 4.3 and 4.4 are devoted to the RDYC and RYC strategy, respectively. Both sections show how the respective strategy is implemented and aim at giving first impressions about the returns obtained from the respective strategy. To achieve this goal, each section contains the excess return-volatility profile of the corresponding strategy. The following subsection 4.5 covers the results of the performance evaluation of both strategies. It starts with subsection 4.5.1, which presents various performance indicators that will be applied in the later parts to assess the strategies and includes traditional, value at risk-based as well as lower partial moment-based performance measures. The following two sections contain the performance results of both strategies relatively to a benchmark strategy, which is the German Stock Index (DAX) for the RDYC strategy and buy-and-hold (BH) strategy for the RYC strategy. Both sections provide the results of unconditional strategies as well as those where a filter rule was applied.

Independently of the business strategy, banks have to control their exposure to various risks. The last subsection of section 4 gives an overview over the main sources of risk as well as current and planned regulatory framework, which aims to maintain the stability of the financial system. Subsection 4.6.1 is related to the Basel II regulations which are currently in use. In addition, new developments regarding the regulatory requirements are presented, which arose during the financial crisis of 2008. Subsection 4.6.2 addresses the specific risks attributable to the yield curve strategies. It also emphasizes the new regulatory requirements of Basel

III related to specific risks of the strategies, such as liquidity risk. Finally, chapter 5 provides a summary of the obtained results.

2. Term Structure of Interest Rates

This chapter is devoted to various theories that were elaborated to explain the term structure of interest rates. Section 2.1 introduces different estimation methods available for deriving the term structure. Sections 2.2 and 2.3 outline the main focus of the classical term structure theories including the pure expectations theory, liquidity preference theory, preferred habitat theory and the market segmentation theory and address their ability to explain various shapes of the yield curve. In this part, several procedures commonly applied to test various term structure theories will be presented. Subsequently, the theoretical foundation is followed by a summary of existing empirical evidence in section 2.4.

2.1 Estimation Methods

The term structure of interest rates represents the relationship between the spot rates, i.e. interest rates for an investment beginning at the time of consideration, and the term to maturity of the investment. The yield curve plots the relationship between the bond yields and their remaining maturities. It should be distinguished from the forward curve, which plots the relationship between the forward rates and their maturities. Forward rates are interest rates on an investment that starts on some future date and lasts a particular number of periods. As the term structure is not directly observable, it has to be estimated. The term structure can be calculated from: 1) zero bonds; 2) coupon-bearing bonds; 3) swap rates. One way to obtain the term structure is to calculate yields to maturity of the default-free zero-coupon bonds. Zero bonds are fixed income securities that do not provide interest payment during the bond's life and whose single cash flow is the repayment of the face value at the end of their maturity. Due to that fact, they allow a straightforward calculation of the spot rates from the observed prices. In order to obtain the price P of a zero-bond with the maturity T, the single cash flow has to be discounted with an appropriate interest rate:

$$P = \frac{\mathrm{FV}}{\left(1 + r_T\right)^T}, \tag{1}$$

where FV represents the face value of the bond. A corresponding interest rate can thus be easily obtained by solving the above equation for r_T. Theoretically, if prices of zero-coupon bonds were available for every required maturity, the term structure could be easily derived. However, most of the traded bonds are coupon-bearing and there are not enough zero bonds available to estimate the whole maturity spectrum.

Spot rates could also be determined as yields to maturity of coupon-bearing bonds of similar credit worthiness. These bonds pay an interest rate, called a coupon rate, every period during the bond's life and are normally available for a brighter maturity spectrum than zero bonds. To derive the spot rates, government bonds are often selected, as they are considered to be free of default risk. The price of a coupon-bearing bond can be determined in the following way:[1]

$$P = \frac{C_1}{(1+r_1)^1} + \frac{C_2}{(1+r_2)^2} + ... + \frac{C_T + \mathrm{FV}}{(1+r_T)^t}, \tag{2}$$

where C denotes the coupon payment. The term structure could be then estimated as yield to maturity (YTM) of coupon bonds. It represents the internal rate of return (IRR) on an investment in a bond, at which the present value of its cash flows equals the price of the bond. YTM is widely used as an indicator of an average rate of return of a bond throughout its life and can be calculated from the following equation:[2]

$$P = \frac{C_1}{(1+y)^1} + \frac{C_2}{(1+y)^2} + ... + \frac{C_T + \mathrm{FV}}{(1+y)^t}, \tag{3}$$

where y denotes the yield to maturity. However, this method has its disadvantages, as bonds of the same maturity but different coupon rates could have different yields. Moreover, calculating YTM assumes that the coupon payments are reinvested at the IRR. Consequently, YTM of a coupon bond will only coincide with the respective spot rate in the case that the interest rates are the same for all maturities.[3] In all other cases yields to maturity of coupon bonds may only approximate the term structure of interest rates. Another method to derive the term structure is to use the *swap rates*. A swap represents a contract in which two parties, usually banks of high credit worthiness, agree to exchange series of cash flows. In an interest rate

[1] See Jarrow/Turnbull (2000), pp. 386-393.
[2] See Bodie/Kane/Marcus (1999), p. 417.
[3] See Fabozzi (2004), pp. 99-100.

swap, the series of interest payments on some principal amount are exchanged. As a rule, fixed interest rate payments are exchanged for floating rate payments. Using the fixed interest rates quoted by the financial institutions known as swap rates, the spot interest rates can be determined. The value of the swap can be determined as the difference between the value of a coupon bond and a bond with floating payments, a floating rate note. At the point of initiation, the value of the swap contract is set to be zero. Consequently, the fixed side of the swap can be viewed as a coupon bond selling at its par value, which is the case when the coupon rate is equal to the yield to maturity. The value of such a bond with the maturity of two years can be expressed as follows:[4]

$$\frac{sr_2}{(1+r_1)^1}+\frac{100+sr_2}{(1+r_2)^2}=100, \tag{4}$$

where sr_2 stands for a swap rate of a two-year swap. If the one-year spot rate r_1 is known, the two-year spot rate r_2 can be easily derived. Similarly, spot rates of longer maturities can be determined:

$$r_m=\left(\frac{100+sr_m}{100-\sum_{i=1}^{m-1}\frac{sr_m}{(1+r_i)^i}}\right)-1, \tag{5}$$

where sr_m stands for a swap rate of an *m*-year swap and r_i stands for an *i*-year spot rate. After the spot rates have been obtained, the continuous term structure can be derived by using *linear interpolation*. In addition to the above methods, the continuous term structure of interest rates can be estimated with the help of theoretical models. The motivation behind these models is to receive a more precise continuous term structure than by means of interpolation methods. The essence of such models is to calculate the theoretical yields to maturity, assuming some particular functional relationship for the term structure of interest rates. The parameters of this function are then estimated in such a way, that the observed bond prices match the cash flows discounted by the theoretical spot rates as close as possible.[5] Some commonly used ex-

[4] See Hull (2005), pp. 149-155.
[5] See Martellini/Priaulet/Priaulet (2010), pp. 117-122.

amples include the spline-based method,[6] the Nelson-Siegel approach[7] and the Svensson approach.[8] In the Nelson-Siegel approach, the interest rate of maturity *n*, is a function of four parameters:

$$r^n = \varphi_0 + \varphi_1 \left(\frac{1 - \exp\left(-\frac{n}{\varsigma}\right)}{\frac{n}{\varsigma}} \right) + \varphi_2 \left(\frac{1 - \exp\left(-\frac{n}{\varsigma}\right)}{\frac{n}{\varsigma}} - \exp\left(-\frac{n}{\varsigma}\right) \right), \tag{6}$$

where r^n is the continuously compounded spot rate with maturity *n*; φ_0 φ_1, φ_2 and ς are the parameters. In the framework of this approach, φ_0 can be interpreted as the long-term level of interest rates; φ_1 corresponds to the slope of the yield curve; φ_2 is a curvature parameter; ς can be identified as a speed of decay of the short- and medium-term rates to zero. The popularity of this method is based on its ability to capture all typical shapes of the term structure and a reasonable number of parameters to be estimated. Moreover, these parameters allow a clear interpretation as level, slope and curvature of the yield curve. The extension of this approach was performed by L. Svensson, who extended the model to five parameters. This allowed capturing nearly all possible shapes of the term structure.[9] The Nelson-Siegel and the Svensson approach are being widely used by central banks.[10]

Historically, several types of shape of the term structure have been observed. They can be generally divided into a normal, flat and inverse term structure, although more exotic forms such as humped or u-shaped are also possible. The normal term structure, as its name already says, is the one which is commonly being observed. The main feature of this type of shape is that the long-term interest rates lie above the short-term interest rates. The yield curve is then upward-sloping, as depicted in figure 1 for the German average spot rates ranging from 1972 to 2007. The flat shape corresponds to the situation when the interest rates are equal, independently of their maturity. Finally, the inverse term structure is characterized by a down-

---

[6] Polinomial splines were introduced by McCulloch (1971). Later, the exponential spline-based method was developed by Vasicek/Fong (1982). The difficulty of this method is in the determination of the optimal number of splines.
[7] See Nelson/Siegel (1987).
[8] See Svensson (1994).
[9] See Martellini/Priaulet/Priaulet (2010), pp. 117-122.
[10] For example, the German Federal Bank uses the Nelson-Siegel and Svensson approach to estimate the term structure, see German Federal Bank (1997).

ward-sloping yield curve, i.e. short-term interest rates exceed the long-term interest rates. Although the inverse and flat term structures occur less frequently than the normal type, they could be observed in Germany over the past 30 years. For example, the inverse term structure, plotted in figure 2, has occurred in September 1981. In turn, figure 3 presents the case of a flat term structure, which could be observed in March 2007.

Figure 1: Average Spot Rate in Germany, 1972 – 2007

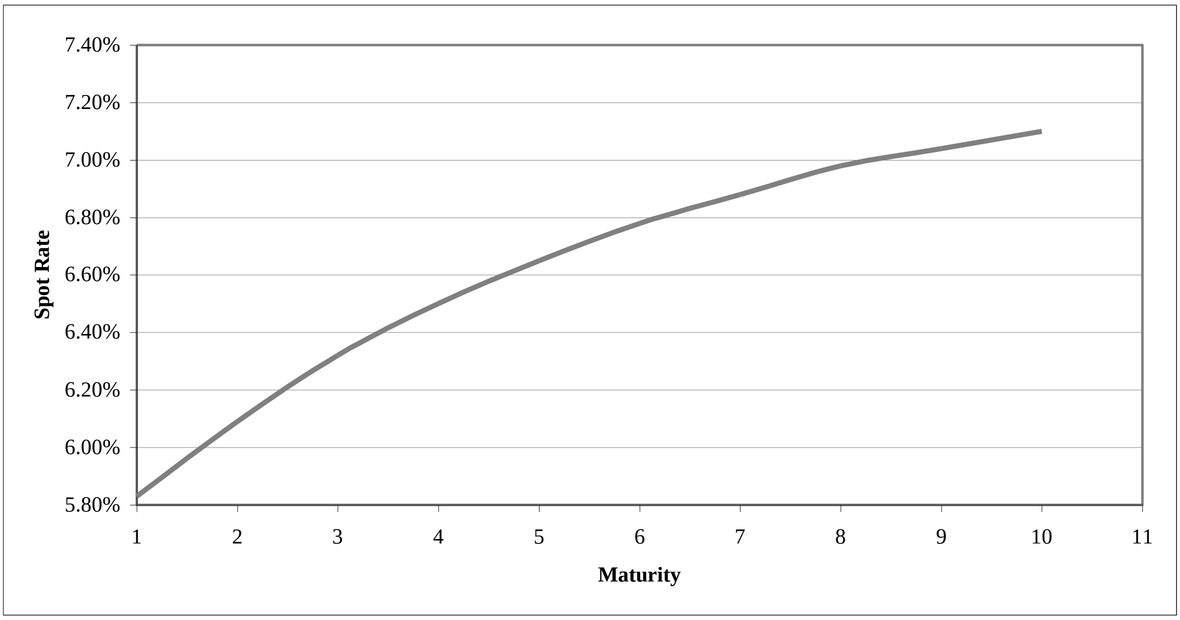


Figure 2: Term Structure of the German Interest Rates, September 1981

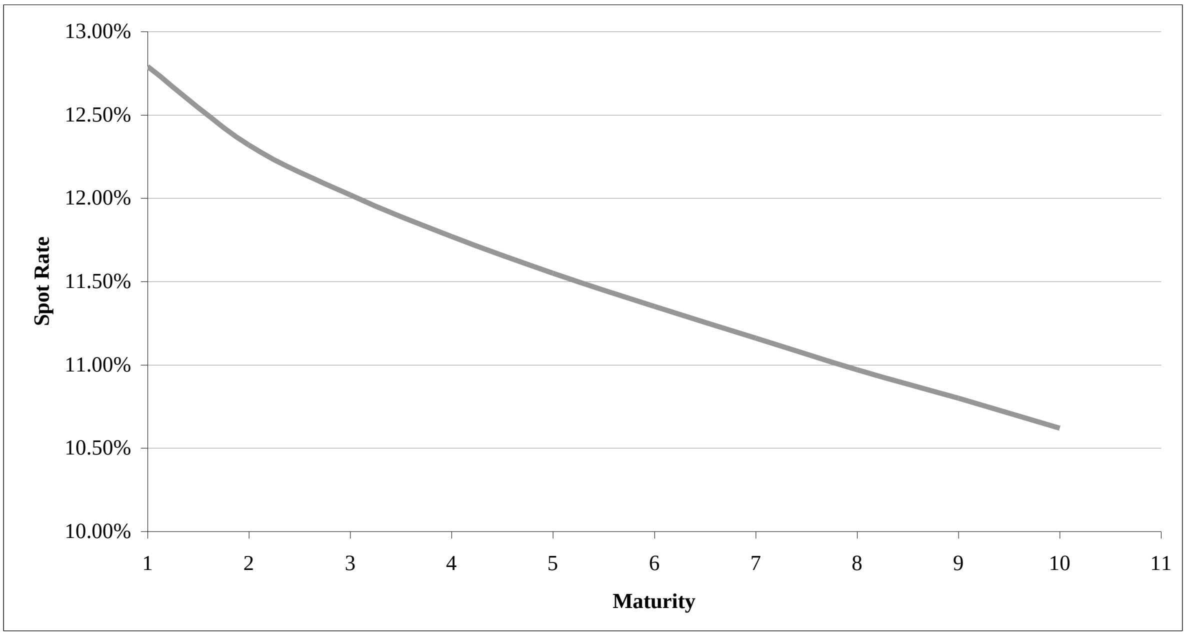
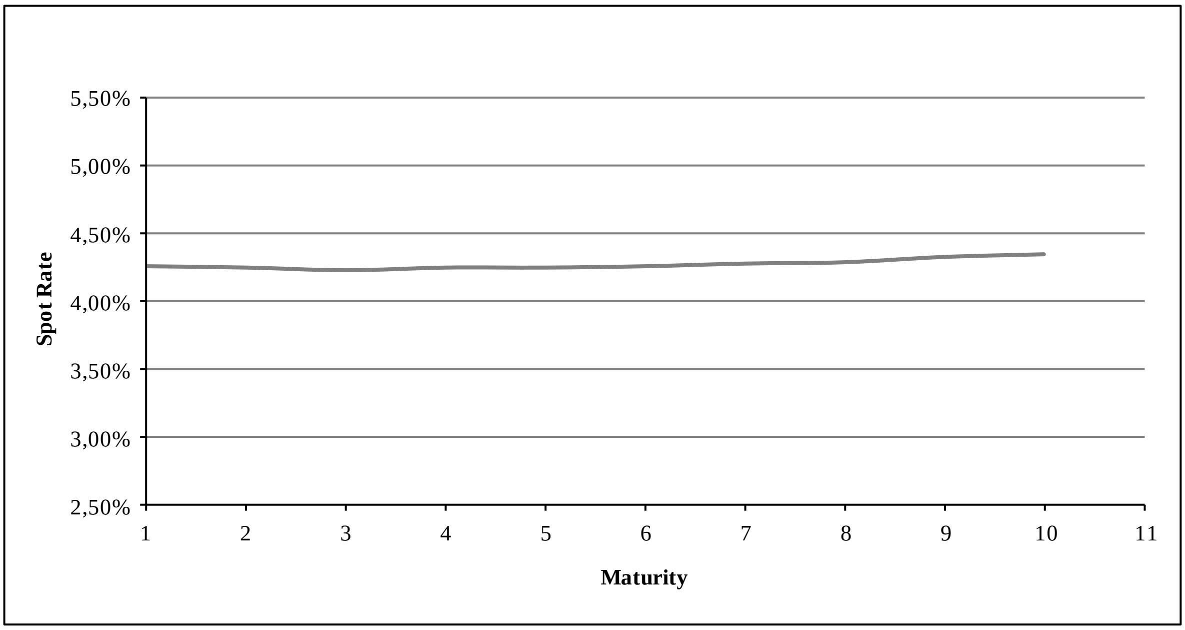


Figure 3: Term Structure of the German Interest Rates, March 2007



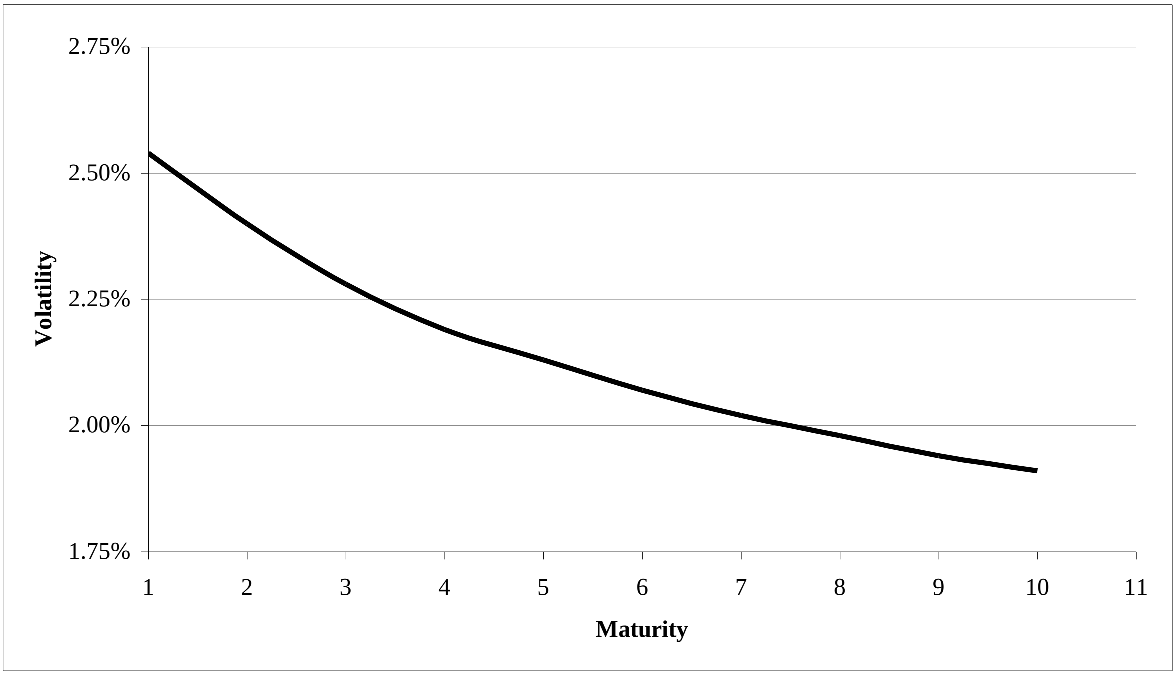
Several further observations connected with the yield curve are worth mentioning and can be summarized as follows:

- The yield curve is upward sloping most of the time. However, the yield curve can be upward-, downward-sloping as well as of nearly zero slope.
- Interest rates of different maturities tend to move together. Short-term interest rates do not change independently of the long-term interest rates. This empirical fact is reflected in figure 4, which plots the one-year interest rate as well as the ten-year interest rate over the period 1972 – 2007.
- Short-term interest rates are more volatile than the long-term interest rates. This fact can be easily seen from figure 5, which plots the volatility of average German interest rates against their maturity for the period 1972 – 2007. The short-term rates exhibit the highest volatility of 2.5 percent, which decreases with the increasing maturity of interest rates. The volatility of the ten-year spot rate constitutes only 1.91 percent.
- Interest rates tend to lie in some range, i.e. they do not rise beyond a certain level.

If interest rates are very high, they usually fall again after reaching some certain level. The same applies to the situation when interest rates are unusually low: they tend to return to some historical normal level. Thus, interest rates stay within some certain range.

Figure 4: German One-Year versus Ten-Year Spot Rate, 1972 – 2007

Figure 5: The Volatility of Spot Rates, 1972 – 2007



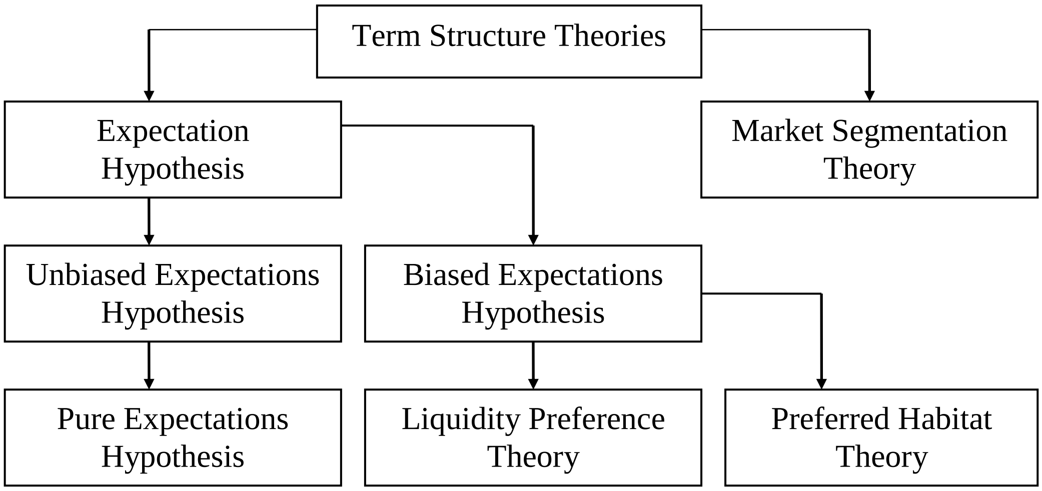
The empirical facts described above gave rise to several term structure theories. Each of them attempts to explain the observed behavior of the term structure of interest rates. Especially the first three empirical facts found a strong reflection in the literature on term structure. In the following the classical term structure theories and different empirical facts related to them will be presented.

2.2 The Expectations Hypothesis of the Term Structure

The expectations theory together with the market segmentation theory belongs to the classical theories of the term structure of interest rates. The EH[11] is a common term used to summarize the term structure theories that explain the long-term interest rates by means of expectations of economic agents. It is comprised of several forms: the pure expectations theory, the liquidity premium theory, and the preferred habitat theory. In its pure version which was originally proposed by Irving Fisher,[12] the expectations theory assumes that the term structure is determined entirely by the expectations of the future short-term interest rates. In contrast to the pure expectations theory, two other forms of the EH state the existence of some additional factors explaining the term structure. Consequently, they are referred to as biased expectations theories. In contrast, the market segmentations theory does not incorporate the expectations of the future interest rates. Instead, it states that interest rates depend on the interaction between the demand and supply.[13] Figure 6 gives an overview of the classical term structure theories.

[11] The terms „expectations theory“ and „expectations hypothesis“ are used interchangeably in this thesis.
[12] See Fisher (1896).
[13] See Mishkin (1994), p. 113.

Figure 6: The Classical Term Structure Theories



2.2.1 The Pure Expectations Hypothesis

Advocates of the pure expectations hypothesis of the term structure state that investors do not have preferences towards bond's maturity. The only selection criterion in their investment choice is the bond's expected return, i.e. the investors are assumed to be risk-neutral. According to this argument, bonds of different maturities are perfect substitutes. Therefore, an investor with a five-year investment horizon will be indifferent between buying a bond with the maturity of five years or rolling over five one-year bonds. Thus, if a bond of some certain maturity has a lower expected rate of return compared to another bond, the former will not be purchased. For example, if the current term structure is flat and the future short-term interest rates are expected to rise, investors will decide to buy a short-term bond rather than a long-term one. After one year, they can reinvest the proceeds at a higher rate, according to their expectations. As a result, the price of a long-term bond will decrease, which will lead to a higher return on that bond. The term structure will be no longer flat, but upward-sloping. Thus, in the framework of the pure expectations hypothesis, the bond's maturity does not play a role.[14]

The pure expectations hypothesis asserts that long-term spot rates are equal to the geometric mean of current and expected future short-term rates. The slope of the term structure thus reflects the current expectations of market participants regarding future short-term rates. If short-term rates are expected to rise, the yield curve will have a positive slope. In case that

[14] See Cuthbertsin/Nitzsche (2004); pp. 494-495.

market participants do not anticipate the short-term rates to change,[15] a flat yield curve will be observed. Finally, a downward-sloping yield curve will indicate that the short-term interest rates are expected to fall. Thus, the pure EH can be indentified with every shape of the yield curve. This version of the expectations theory has been subject to intensive testing. In order to enlighten these testing procedures, it is convenient to state the pure expectations theory in mathematical terms. The pure EH can be stated as:[16]

$$r_t^n = \left[\left(1+r_t^{0,1}\right)\cdot\left(1+E_t r_{t+1}^{1,2}\right)\cdot\left(1+E_t r_{t+2}^{2,3}\right)\ldots\left(1+E_t r_{t+n-1}^{n-1,n}\right)\right]^{1/n} - 1, \tag{7}$$

where r_t^n is the rate of return on a bond with maturity n and t refers to the time today. Expectation terms denote the expectations of future one-year short-term rates on an investment starting in $t+i$, i=0,1,…,n periods from now. The pure expectations theory states that forward rates fully reflect the expected future interest rates:

$$f_t^{t+i,t+i+m} = E_t r_{t+i}^m, \tag{8}$$

where $f_t^{t+i,t+i+m}$ stands for the forward rate determined today for a contract starting in $t+i$ and ending in $t+i+m$ and r_{t+i}^m is the future spot rate for a contract starting in $t+i$ which lasts for m periods. Equation (8) implies forward rates to be unbiased predictors of future spot rates. In the context of monetary policy, equation (7) would mean that the only possibility to affect the long-term interest rates is to influence the expectations of market participants. Changing the short-term interest rate without influencing the expectations would lead to an insignificant influence on the long-term interest rate. For example, if policy makers increase the one-month interest rate by 100 basis points (bps) and this change is expected to be temporary, the interest rate on a 10-year bond will only increase by approximately 100 bps/120, i.e. by less than one basis point.[17] Only if the change in the short-term rate is expected to be of a permanent nature, will the long-term rates rise by 100 basis points. From equation (7), the three-year spot rate is given as:

[15] Such behavior of economic agents is sometimes called “static expectations”.
[16] See Walsh (2003), pp. 491-492.
[17] See Sorensen/Whitta-Jacobsen (2005), p. 511.

$$r_t^3 = \left[\left(1 + r_t^{0,1}\right) \cdot \left(1 + E_t r_{t+1}^{1,2}\right) \cdot \left(1 + E_t r_{t+2}^{2,3}\right)\right] - 1. \tag{9}$$

From the above equation, the implicit expectations of market participants about future expected short-term rates can be derived:

$$E_t r_{t+2}^{2,3} = \frac{\left(1 + r_t^3\right)^3}{\left(1 + r_t^{0,1}\right) \cdot \left(1 + E_t r_{t+1}^{1,2}\right)}, \tag{10}$$

In the empirical literature, a linearized version[18] of formula (7) is widely applied:

$$r_t^n = \frac{1}{n} \sum_{i=0}^{n-1} E_t r_{t+i}^1 + \upsilon^{n,1}. \tag{11}$$

Equation (11) states that the *n*-period interest rate is explained by the simple average of the current and future expected one-period[19] interest rates plus a constant risk or term premium. Equation (11) represents a weaker version of the EH, as in the pure expectations hypothesis the term $\upsilon^{n,1}$ is equal to zero. The next step is to subtract the term r_t^1 from both sides of the equation (11) and rearrange the terms to receive the following:

$$r_t^n - r_t^1 = \sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right) E_t \left(\Delta r_{t+i}^1\right) + v^n. \tag{12}$$

According to equation (12), the spread between long-term and short-term interest rates can be explained by the difference in the expected future one-period interest rates plus a term premium. As expectations of market participants are not known, a typical assumption is that expectations are formed rationally:[20]

$$E_t r_{t+i}^1 = r_{t+i}^1 + \varphi_{t+i}. \tag{13}$$

[18] Under the approximation $\ln(1+r) \approx r$.
[19] This expression can also be generalized to *m*-period short rates.
[20] The rational expectations hypothesis was originally proposed by J. Muth in 1961 in his paper "Rational Expectations and the Theory of Price Movements". In the 1970-s, the hypothesis was further developed by Robert Lucas (1972) and Thomas Sargent (1973). The rational expectations hypothesis states that individuals are able to make correct predictions of macroeconomic variables based on all available information in the period where the forecast is made. Although individuals do not possess deep knowledge of complex economic models, they can use forecasts produced by the professionals and thus make best possible predictions which can be made conditional on the available information set.

With equation (13) a testable version of the EH is obtained:[21]

$$r_t^n - r_t^1 = \alpha + \beta\left(r_{t+i}^1 - r_t^1\right) + \eta_t. \tag{14}$$

In this framework, the pure EH is tested by estimating equation (14) and testing the null hypotheses α =0 and β =1. If the null is rejected but significance of β is confirmed, this result is usually interpreted as an evidence of forward rates having explanatory power. The above equation is an example of a regression that predicts changes in the short-term rate. Alternatively, it can be tested, whether the term spread can forecast changes in the long-term interest rate. This is done with the help of the following equation:

$$r_t^n - r_t^1 = \alpha + \beta\left(r_{t+1}^{n-1} - r_t^n\right) + \eta_t. \tag{15}$$

For testing purposes the forward-spot spread approach is also frequently adopted. This approach is similar to equation (14); the only difference constitutes the term in brackets on the right hand side. Instead of the difference between the future short-term rates, the forward-spot spread is applied:

$$r_t^n - r_t^1 = \alpha + \beta\left(f_t^1 - r_t^1\right) + \eta_t. \tag{16}$$

According to this formulation, the spread between the long- and short-term spot rates can be explained by the forward-spot spread. Using formula (8), it can also be directly tested if forward rates can predict future spot rates. Then the null hypotheses are: α =0, β =1 for the pure EH and β =1 for the biased expectations theory:

$$r_{t+i}^m = \alpha + \beta_1 f_t^{t+i,t+i+m} + \phi_t. \tag{17}$$

The pure expectations theory provides an explanation why interest rates on bonds with different maturities move together over time. As long-term interest rates merely represent the average of expected future short-term rates, a rise in short-term interest rates will cause long-term

[21] For discussion of different testing equations for the pure expectations hypothesis, see Culbertson (2004), pp. 520-523.

interest rates to rise as well. The statement that long-term interest rates express the average of expected future short rates also implies that it is not profitable to borrow at the short-term rate and buy long-term bonds, even though the long-term interest rate may lie above the short-term rate.

The pure expectations hypothesis is appealing due to its simplicity and ability to fit every typical shape of the term structure. In addition, as long-term interest rates represent a weighted average of the current and expected short-term interest rates, the long-term rates should be less volatile than the short-term interest rates. Thus, the second empirical fact associated with the yield curve is explained by the pure expectations theory. The same line of reasoning leads to the explanation of the interest rates moving together over time. However, the pure EH fails to explain an important empirical observation related to the behavior of interest rates, namely, that typically an upward-sloping yield curve is observed. Such a behavior of interest rates is inconsistent with this theory, as it would imply that investors expect rising interest rates most of the time. However, rising as well as declining interest rates may occur.[22]

The pure expectations theory ignores risks associated with an investment in bonds. If forward interest rates would perfectly predict future spot rates, there were no uncertainty with respect to future bond prices. However, in reality future bond prices are not known. Thus, an investor with a five-year investment horizon could buy a bond with a maturity of five years or, for example, a ten-year bond and sell it after five years. However, he is uncertain about the price of a ten-year bond that will prevail five years from now. Put differently, the bonds are subject to price risk and, in addition, reinvestment risk. Consequently, there was a need to elaborate a theory that could explain why the yield curve is upward-sloping most of the time. Such a theory is known as the liquidity premium theory.

[22] See Mishkin (1992), pp. 144-145.

2.2.2 The Liquidity Preference Hypothesis

The liquidity preference theory developed by Hicks in 1946[23] emphasizes an uncertainty connected with long-term securities, which are subject to inflation risk and interest rate risk. As a compensation for this uncertainty, market participants demand a positive liquidity premium for holding a longer-term security. The liquidity preference theory states that the shape of the yield curve is determined by two factors: the expectations of future interest rates and a premium for holding a long-term bond, known as liquidity premium. Forward rates implied by the term structure are, therefore, no longer unbiased predictors of the future short-term rates as, in addition to the expectations of the future short-term rates, they contain a liquidity premium, which increases with the time to maturity. Under the liquidity preference hypothesis, forward rates can be expressed as:

$$f_t^{t+i,t+i+m} = E_t r_{t+i}^m + L^m, \tag{18}$$

where L^m is a liquidity premium for holding a bond of maturity *m*. Tests of the liquidity premium theory are based on the following equation:

$$r_t^n - r_t^1 = \alpha + \beta\left(r_{t+i}^1 - r_t^1\right) + \eta_t. \tag{19}$$

Then the null hypothesis is α >0, β =1. The intercept term α is interpreted as the liquidity premium, i.e. the amount by which the long-term interest rate exceeds the expectation of the future short-term rates. If, instead, the forward-spot spread approach is taken, the test equation looks as follows:

$$r_t^n - r_t^1 = \alpha + \beta\left(f_t^1 - r_t^1\right) + \eta_t. \tag{20}$$

Then the null hypothesis is, as before, α >0, β =1. Different variations of the liquidity preference theory can be distinguished through the assumption of a constant or time-varying liquidity premium. The former states that, although the liquidity premium increases with the term to maturity, it stays constant over time. In contrast, advocates of the time-varying liquidity pre-

[23] See Hicks (1946), pp. 141-145.

mium argue that it should rise with the bond's maturity and, in addition, vary over time. The forward rate will be then given as:

$$f_t^{t+i,t+i+m} = E_t r_{t+i}^m + L_t^m \tag{21}$$

where L_t^m is the liquidity premium for holding a bond of maturity *m* determined at time *t*. Similarly to the pure expectations theory, the liquidity preference hypothesis views long-term rates as average expected short-term rates and therefore provides an explanation of the short-term and long-term interest rates moving together. Despite the fact that economic agents may have any kind of expectations towards the future short-term rates, an upward-sloping yield curve will be the most common because of a positive liquidity premium, increasing with time to maturity. Moreover, this version of the expectations theory fits to every shape of the yield curve. For example, even if expected short-term rates are falling, a presence of a liquidity premium that rises with the maturity may result in an upward-sloping yield curve. Likewise, a combination of a declining expected short-term rate and a constant liquidity premium may yield a downward-sloping term structure. The resulting yield curves are a combination of expectations about the future short-term rates and either constant or variable liquidity premium.

2.2.3 The Preferred Habitat Theory

Likewise, the preferred habitat theory, usually associated with the work of Modigliani and Sutch (1966), asserts that the yield curve is formed by the expectations of future short-term interest rates as well as a risk premium. The main difference from the liquidity preference theory is that this premium does not rise uniformly with the instrument's maturity. As opposed to the liquidity preference theory, within this version of the expectations hypothesis, investors do not necessarily prefer shorter-term securities. Instead, investors have different preferred investment horizons or habitats. Thus, if supply and demand for a given maturity range do not match, a risk premium is required to induce market participants to buy bonds outside their maturity preference or habitat. The risk premium can be either positive or negative.[24]

[24] See Mishkin (1992), pp. 154-155.

Due to its expectations component, the preferred habitat theory provides an explanation of yields on bonds of different maturities moving together. The fact that the yield curve is usually upward-sloping is also explained by this modification of the expectations theory. Most investors have a short habitat and, therefore, require a premium for holding longer-term securities. Thus, even if short-term interest rates are not expected to change in the future, long-term interest rates exceed the short-term interest rates. The preferred habitat theory captures all possible shapes of the term structure. Although a premium is positive, a downward-sloping yield curve can arise in case market participants expect a dramatic decline in short-term interest rates. Then, even considering a positive premium, the average of the future expected short-term rates, i.e. the long-term rates, will still lie below the short-term interest rate.[25]

The preferred habitat theory also enables an easy interpretation of the investors' expectations from observing the yield curve. A flat curve would arise as a combination of falling expected interest rates and a positive premium. An upward-sloping, but not very steep yield curve would imply static expectations regarding the future short-term interest rates. Finally, a very steep upward-sloping curve would indicate that market participants anticipate rising interest rates.[26]

2.3 The Market Segmentation Theory

The most well-known alternative to the expectations theory is the market segmentation theory developed by Culbertson (1957).[27] It states that investors have particular holding periods that they strongly prefer and, therefore, only choose bonds that match their investment horizons. As an example of an investor with a short-term horizon commercial banks are usually mentioned whereas pension funds are said to have a long-term investment horizon. Consequently, the interest rate on bonds of different maturities results from the interaction of demand and supply for the respective bond. Markets for short-term and long-term bonds are considered to be completely separated from each other. Whereas the pure EH considers bonds of different maturities to be perfect substitutes, advocates of the market segmentation theory do not regard

[25] See Mishkin (1992), pp. 144-145.
[26] See Mishkin (1992), p. 146.
[27] See Culbertson (1957), pp. 489-504.

them as being substitutes at all. The demand for bonds of a particular maturity is not affected by the expected return on bonds of another maturity.[28]

Under the market segmentation theory, different shapes of the yield curve are the result of the demand for bonds of a particular maturity prevailing at that time. Thus, an upward-sloping yield curve is the most common because usually the demand for short-term bonds dominates. Consequently, such bonds have a higher price and lower interest rate than long-term bonds. However, as markets for bonds of different maturities are completely segmented, there should be no reason for yields on short-term and long-term bonds to move together. This contradicts the empirical observation that bonds of different maturities tend to move together and is the main shortcoming of the market segmentation theory.

Thus, different term structure theories perform differently with respect to empirical facts that were observed in connection with the yield curve. Among them, the liquidity preference theory and, very similar to it, the preferred habitat theory seem to be consistent with all three empirical observations. Table 1 summarizes the classical term structure theories and addresses their ability to explain the empirical facts associated with the yield curve. The question which of the theories is also consistent with the empirical data gave rise to numerous empirical papers. Thus, all three versions of the expectations theory as well as the market segmentation theory were subjects to extensive testing. Although it is barely impossible to cover the whole empirical research on the subject, the next section attempts to reflect the main results.

2.4 Previous Empirical Findings

First formulations of the expectations theory appeared already in the end of the 19th century. However, the theory was fully developed only in the 30s of the past century. First empirical tests of the expectations theory date back to the 1970s. They were conducted using US data and employed simple regression techniques. Since that, a great variety of tests has been performed which examined different implications of the theory, using different methods and maturities. A great majority of these studies, however, concentrates on the US data. The early literature on the term structure can be divided into two categories: studies that use the term spread and studies that apply the forward spot rate for testing the EH. Those preferring the term spread usually perform the test in both directions. In addition, these studies can be dis-

[28] See Martellini/Priaulet/Priaulet (2010), p. 85.

tinguished according to the applied technique. While early literature predominantly applied linear regression techniques, researchers switched to more sophisticated methods in the 1990s. Table 2 at the end of this section provides a summary of the selected studies.

Table 1: Term Structure Theories: Main Characteristics

	Empirical fact	Pure Expectations Theory	Liquidity Preference Theory	Preferred Habitat Theory	Market Segmentation Theory
I Main characteristics					
1	Relation between short-term and long-term rates	Long rates are average expected future short rates	Long rates are average expected future short rates plus a premium	Long rates are average expected future short rates plus a premium	Long-term and short-term interest rates are not related to each other
2	Premium	No premium	Premium for holding long-term securities	Premium for the deviation from preferred habitat	No premium
3	Degree of substitution between short-term and long-term bonds	Perfect substitutes	Imperfect substitutes	Imperfect substitutes	No substitution possible
II Ability to explain empirical facts					
1	Interest rates of different maturities move together	Explains	Explains	Explains	No adequate explanation
2	Short-term interest rates are more volatile than the long-term interest rates	Explains	Explains	Explains	No adequate explanation
3	Yield curve is typically upward-sloping	No adequate explanation	Explains	Explains	Explains
4	Yield curve may be upward-, downward-sloping or have a zero slope	Explains	Explains, if extremely low future interest rates are expected	Explains	Explains

Early studies for the US undoubtedly reject the pure EH and find poor explanatory power of forward rates as well as term spreads. Among them are Hamburger and Platt (1976), Fama (1976), Shiller, Campbell and Schoenholz (1983) and many others. Fama (1984), who investigated short-term interest rates for the period 1959 – 1982, although rejects the pure EH, suggests some predictive power of the forward-spot spread towards the spot rate one month ahead. Later, Mishkin (1988) confirms these findings using a slightly longer period, 1959 – 1986. According to his results, forward-spot spreads can predict changes in the short-term interest rate up to three month in advance.

While previous studies focused on the maturities below one year, Fama and Bliss (1987) analyze the information content of a one-year forward rate from 1964 to 1985. According to their study, there is little predictive power of the forward-spot spread on the short-term forecasting horizon. However, long-term forward rates exhibit significant predictive power for longer forecasting horizons which, according to the authors, can be explained by a slow mean reversion of spot rates. Jorion and Mishkin (1991) also use the forward-spot spread approach to forecast changes in one-year interest rates over the period 1973 – 1989. As opposed to the findings of Fama and Bliss (1987), they conclude that the information content of the spread is poor in the US data, as little predictive ability was found both on the short- and the long-term horizon. Only in the case of Germany and Switzerland, they were able to confirm the predictive power on a five-year horizon.

Regressions employing term spread also did not yield uniform results. Mankiw and Summers (1984) could not confirm the ability of the term spread to forecast changes in the US short-term interest rates from 1963 to 1983. Mankiw and Miron (1986) use a long sample of three- and six-month interest rates ranging from 1890 to 1979 to test whether the slope of the yield curve may be useful for predicting changes in the spot rates. Whereas their study documents little predictive power of the spread for the period after 1915, the year in which the Federal Reserve was founded, the EH proves to be consistent with the data before 1915. The authors attribute poor performance of the expectations theory after 1915 to the increased role of the Federal Reserve System. The interest rate stabilization policy, conducted by the Federal Reserve, could cause a random walk behavior of the short-term interest rates and, therefore, be a reason for the earlier failures of the expectations hypothesis. Later, Hsu and Kugler (1997) find significant support for the predictive power of the term spread towards the changes in the short-term rate in the US over the period 1987 – 1995. However, prior to 1987, the predictive power is poor. Similarly to the study of Mankiw and Miron (1986), the authors attribute this result to the actions of the Federal Reserve, who conducted monetary policy dependent on the term spread starting from the late 80s.

A number of studies document a so called “sign puzzle” in regressions that use term spreads for forecasting changes in long-term interest rates. The essence of this puzzle is that the term spread predicts the wrong direction of the long rate dynamics, i.e. when the difference between the long-term and short-term rates is positive, a decline in long-term interest rates is predicted. The sign puzzle received significant attention in the term structure literature. Among the authors investigating the issue are Campbell and Shiller (1987), Fama (1984), Fama and Bliss (1987), Mishkin (1988). In a more recent study Campbell and Shiller (1991) adopt the vector autoregression (VAR) approach to test the EH with the yields on US treasury bills for a variety of maturities. They assert that the term spread only has significant forecasting ability with respect to changes in short-term, but not in long-term spot rates.

The poor support of the EH by the empirical data caused numerous attempts to further develop the theory which would be able to address this failure. The development of new econometric techniques such as cointegration[29] gave rise to a new wave of research in the area. However, the mixed character of the early results for the expectations theory in US data persisted. Whereas Engsted and Tanggaard (1994) document cointegration in the US term structure for the period 1952 – 1987, Mustafa and Rahman (1995), who examined almost the same period, found no relationship between the long-term and short-term interest rates in the quarterly data ranging from 1953 to 1992. In contrast, Nourzad and Grennier (1995) found forward rates and spot rates in the period from 1981 to 1994 to be cointegrated.

Despite the fact that some of the above studies confirm some explanatory power in forward rates or short-long spreads, they, in general, statistically reject the pure EH. Tests of the expectations hypothesis considering European data are, basically, more supportive to the expectations theory of the term structure.

[29] A presence of cointegration would imply the existence of a long-run relationship among the interest rates and could be viewed as an evidence in favor of the EH. The details of this method will be enlightened later in section 3.2.3.

For Germany, there have been only a few studies. Kugler (1988) analyzes the influence of monetary policy on the predictive ability of the term spread using the German, US, and Swiss three- and six-month interest rates over the period 1974 – 1986. For the US data, no predictive power of the term spread towards changes in short-term interest rates was found. In contrast, this study confirms substantial predictive power in case of Germany and Switzerland. In the period of investigation, the Federal Reserve used an interest rate stabilization policy whereas monetary authorities of Germany and Switzerland committed themselves to money supply targeting. The authors interpret this result as evidence in favor of the hypothesis found in Mankiw and Miron (1986) that the reasons of poor performance of the expectations hypothesis lie in the interest rate targeting policy.

Hardouvelis (1994) analyzed the ability of the term spread to predict changes in both long-term and short-term rates for a variety of countries including Germany, Italy, France, USA, Canada, and Japan for the period 1953 – 1992. Although the pure EH is rejected, he reports significant coefficients for the short-term spot rate model for all countries with the exception of Germany and the US. However, when regressing the change in the long-term spot rate on the term spread, his study finds little forecasting ability and documents negative slope coefficients for all countries except of Italy and France. Thus, his results support the "sign puzzle" received in previous studies. The negative slope coefficient, however, disappears if instrumental variables[30] are introduced for all countries with the exception of the US.

Gischer (1996) examined the German term structure for the period 1986 – 1995. He uses forward rates of six preceding periods in order to explain the corresponding spot rate. This study, although finds forward rates to be significant as an explanatory variable, does not confirm the predictive ability of the forward rates towards the one-year spot rate.

Gerlach and Smets (1997) test the predictive power of the term spread with respect to changes in short-term rates for 17 countries including Germany. Their study, considering three-, six- and twelve-month interest rates, provides quite striking results which are considerably in favor of the pure EH. In almost 70 percent of all regressions the null hypothesis that the beta coefficient equals one cannot be rejected. Moreover, in 50 percent of all cases even the joint

[30] Instrumental variables are helpful in removing the correlation between an explanatory variable and the error term in a regression equation. In such cases, the correlation problem can be solved by finding a proxy, called an instrumental variable, which is highly correlated with the explanatory variable, but uncorrelated with the error term. However, to find such a proxy may be a difficult task in practice (see Thomas (1997), pp. 220-221).

hypothesis α =0, β =1 cannot be rejected, which would imply the validity of the pure EH. For Germany over the period 1972 – 1993, the validity of the pure expectations hypothesis cannot be rejected for six- and twelve-month interest rates. This is by far the most supportive result for the pure EH.
Remarkably, only in the case of the US the null α =0, β =1 is rejected for the whole maturity spectrum under consideration.

In contrast, the study of Jondeau and Ricart (1999), who applied both the term spread and the forward-spot spread approach to German, French, UK and US data with maturities less than one year, could not provide such a strong support of the theory. In general, their study for 1975 to 1997 rejected the pure EH for Germany and the US. Moreover, in the regression of forecasting changes in the long-term spot rate, negative slope coefficients were obtained for both countries. In contrast, the EH is generally supported by French and UK data, as β =1 could not be rejected.

Boero and Torricelli (2002) use the estimated German term structure data for 1983 to 1994. They report that the long-short spreads as well as the forward-spot spreads are good predictors for the future short-term spot rates. In contrast, term spreads show little forecasting power with respect to future changes in long-term spot rates. The latter result is consistent with previous findings for the US. However, although the information content is poor, in German data at least the direction of changes in long-term spot rates can be predicted.

The study of Dominguez and Novales (2002) is of particular interest, as the authors examine the ability of forward rates to predict future spot rates for a variety of interest rates using data in levels and not the spreads. They analyze one-, three-, six- and twelve-month interest rates for a variety of countries ranging from Germany to US and Japan for the period 1978 – 1997. Not only they present evidence that forward rates can explain future spot rates to a significant extent, but also the unbiasedness of forward rates cannot be rejected. In addition, the authors investigated the forecasting performance of forward rates, using the estimated coefficients to build predictions for 1998. The study indicates that forward rates can predict spot rates better than can be achieved by using the past values of the spot rates themselves, at least at the short-term horizon.

The popularity of the EH gave rise to research outside the US, UK and European borders. As to this point of time more advanced techniques were available, these studies mostly employ cointegration methods. Guest and McLean (1998) received conflicting evidence about cointegration between the Australian short-term and long-term interest rates and therefore cannot confirm the existence of the long-run relationship in their data. Gonzalez, Spencer and Walz (1999) investigate the relationship between the spot rates and forward rates with the maturities of one, three and six months in Mexico from 1991 to 1996. Their results report significant ability of the forward-spot spread to predict the future short-term interest rates. Cooray (2003) considers three- and six-month spot rates together with the respective forward rates for the case of Sri-Lanka. His results, although rejecting the hypothesis that forward rates are unbiased predictors of the future short-term rates, suggest the existence of cointegration between spot and forward rates. Finally, Tabak (2009) uses Brazilian swap rates for one, three, six and 12 months covering the period 1995 – 2006 to test the expectations hypothesis using term spreads. Although this study rejects the pure EH, the results of the cointegration analysis indicate that the long-short spread is a biased predictor of changes in the short-term interest rates. A presence of a time-varying risk-premium is provided as an explanation for this result.

As described above, in general, the pure EH and its biased versions were not confirmed by empirical papers. The pure version of the EH was rejected by the great majority of the studies. Some of them, however, report some predictive ability of the term spread or the forward-spot spread. This result is especially pronounced for the US and UK data. Evidence for Europe provides more support for the EH. The usage of different time periods for different countries and maturities has resulted in a variety of contradicting findings. Also the data characteristics selected for the test of a term structure theory differ greatly. Some studies apply real data whereas the others apply interest rates data estimated with the help of statistical techniques; this might be a reason for such divergent results regarding the EH. As strong evidence supporting the pure EH could not be found, this gave rise to further research. Many authors address one difficulty connected with the interpretation of test results, namely, the necessity to assume some particular expectation formation process. Thus, it is only possible to check the validity of the expectations theory as a joint hypothesis. Consequently, negative test results can be interpreted in two ways: either the EH does not hold or the expectations are formed in a different way.[31]

[31] Several authors attempted to check the validity of the expectations hypothesis using adaptive expectations, which assume individuals make forecasts based on the past values of economics variables. In this case individu-

Further hypotheses addressing the general failure of the EH include the overreaction hypothesis, presence of measurement errors and time-varying risk-premia. The essence of the overreaction hypothesis is that long-term spot rates over- or under-react regarding the expectations of future short-term spot rates. Mankiw and Miron (1986) attribute the inability of the expectations theory to reliably predict future spot rates to the existence of a time-varying term premium. However, Taylor (1992) could not report any evidence in favor of the pure EH or time-varying premium in the UK data. Instead, he finds support for the validity of the market segmentation theory.

The expectations theory as well as the market segmentation theory represent the oldest theories of the term structure of interest rates. Recently, there have been attempts to develop new methods to explain the term structure. These approaches can be roughly divided into those coming from the financial literature and those originating from the macroeconomic literature. The former model the term structure using pure statistical methods. They are referred to as affine or linear term structure models. In the affine term structure framework, the yield curve can be represented by means of three parameters: the level, the slope and the curvature, which are latent, unobservable factors and do not possess an economic interpretation.[32] Among the study elaborating such a type of model are Longstaff and Swartz (1992), Chen and Scott (1993) and Dai and Singleton (2000). Generally, such models show much better performance in explaining the term structure than the classical theories, as they are able to explain all kinds of movements of the yield curve.

Although the affine term structure models explain the term structure quite well, they do not provide any insight into the connection between the term structure and macroeconomic factors. Consequently, another stand of literature has attempted to connect the term structure with macroeconomic fundamental factors. More recent pure macroeconomic models attempt to explain long-term interest rates not only by means of short-term rates, but also with the help of non-interest variables, such as inflation, exchange rate, business cycle indicators, government borrowing. De Butter and Jansen (2004) find that the German long-term interest rates

als do not only take into account the values observed in the previous period, but possess a “memory”, i.e. also consider the values that occurred in several periods before.

[32] Littermann/Scheinkman (1991) indicate that around 99 percent of all movements of the yield curve can be explained by these factors. Changes in the level happen when interest rates of all maturities rise by approximately the same amount; changes un the slope appear in the case that short-term rates rise at a greater extend than the long-term interest rates; finally, changes in curvature happen in the case that medium-term interest rates rise greater than the short- and long-term interest rates, which leads to a more humped yield curve.

over 1982 – 2001 can be best explained by the German short-term interest rates, foreign long-term interest rates as well as macroeconomic factors such as oil price and economic activity indicators. Wu (2001) employs a VAR model to examine the impact of monetary policy shocks on different parameters of the yield curve in the US in the period 1983 – 1998. His findings indicate that monetary policy shocks mostly affect the slope, but not the level of the yield curve.

The most recent development in the term structure literature represents a mixture of the macroeconomic and finance approach, as it combines latent factors with various macroeconomic factors in order to explain the term structure. The motivation for such a combined approach was a common opinion that, especially at the short-term end of the term structure, interest rates are largely driven by macroeconomic parameters. Ang and Piazessi (2003) find a confirmation of this view using latent factors as well as macroeconomic factors such as inflation rates and different indicators of economic activity in the US. Their findings show that at the short-term end of the yield curve, macroeconomic factors are able to explain around 85 percent of the variation in the interest rates. In contrast, long-term interest rates could be better forecasted by unobserved rather than macroeconomic factors. Consequently, the macroeconomic factors affect the slope and the curvature, but not the level parameter of the yield curve. Using the VAR approach on maturities from one month to one year, Evans and Marshall (2002), however, found that macroeconomic factors not only affect the short- and medium-term interest rate, but also account for around 90 percent of variation in the US long-term rate. They indicate that changes in the level as well as slope and curvature of the yield curve are attributable to such factors.

Hördahl, Tristani and Vestin (2006) suggest that macroeconomic factors had significant impact on the German term structure in the period 1975 – 1998. Their findings indicate that monetary policy shocks mostly influence shorter maturities while inflation and output shocks affect the curvature of the term structure at the medium- and long-term end. In addition, they attest better out-of-sample forecasting performance to the combined models as compared to pure affine term structure models. Table 2 provides a brief summary of the major studies devoted to the pure EH.

Table 2: Summary of the studies devoted to the pure EH

Method	Study	Country/Period	Result
Forward-spot spread	Fama (1976); Hamburger/Platt (1976); Shiller et al. (1983)	US; prior to 1982	Reject the pure EH; no predictive power of the spread
	Fama (1984)	US; 1959 –1982	Rejects the pure EH; some evidence of predictive power
	Fama/Bliss (1987)	US; 1964–1985	Pure EH rejected; some predictive ability of long-term forward rates
	Jorion/Mishkin (1991)	US; 1974–1986	Reject the pure EH; poor predictive power for the US; for Germany predictive power on a five-year horizon
Term spread	Mankiw/Summers (1984)	US; 1963–1983	Pure EH rejected; no predictive power
	Mankiw/Miron (1986)	US; 1890–1979	The data is consistent with the pure EH prior to 1915; no predictive power after 1915 (increased role of the Fed)
	Campbell/Shiller (1987)	US; 1970–1987	The spread predicts only changes in the short- but not in the long-term rate; for the long-term rate, a wrong direction is predicted by the model (“sign puzzle”)
	Kugler (1988)	US/Germany; 1974–1986	Pure EH rejected; the spread possess explanatory power for Germany, but not for the US, explained through the interest rate stabilization policy of the Fed
	Hardouvelis (1994)	US/Germany; 1953–1992	Rejects the pure EH and finds no explanatory power of the spread; for the prediction of the long-term rate, a “sign puzzle” is reported
	Gerlach/Smets (1997)	17 countries including Germany; 1972–1993	In 50 percent of all cases including Germany, the pure EH is supported by the data; in 70 percent of all cases significant predictive power was reported
	Jondeau/Ricart (1999)	US/Germany; 1975–1997	The pure EH is rejected for both counties; “sign puzzle” was found
Forward rates in levels	Gischer (1996)	Germany; 1986–2005	Significant predictive power of forward rates; however poor performance in actual forecasting
	Dominguez/Novales (2002)	US/Germany; 1978–1998	Supportive evidence for the pure EH; confirm the ability to actually forecast the spot rate

3. Testing the Predictive Power of Forward Rates[33]

The goal of this chapter is not to once more test the EH in any of its forms described in the previous section, but to examine whether information helping to predict future interest rates can be extracted from forward rates. Regarding the above mentioned difficulties, the expectations theory will be tested in an indirect way. In particular, it will be determined whether forward rates from past periods, which reflect expectations of market participants in the respective periods, can be used to predict future short-term spot rates. Section 3.1 explains the selected models and introduces the employed econometric techniques. Section 3.2 contains the results of the cointegration analysis and addresses the forecasting ability of forward rates.

3.1 Econometric Methodology

The choice of an appropriate econometric procedure strongly depends on data properties. As the empirical literature mainly applies spreads to test the EH, the problem of non-stationary data was not so pronounced. If spreads are employed, standard regression could be applied for estimating regression coefficients. For this chapter, which uses data in levels and not the spreads, it is important to examine time series properties of the data before deciding on the most appropriate econometric method.

3.1.1 The Model

The analysis in this section aims at examining how well forward rates can predict future short-term spot rates. Thus, as a first step we test equation (17). In addition, it will be checked if forward rates lying farther in the past contain any explanatory power with respect to future spot rates.[34] In other words, if r_t^1 is today's one-year spot rate then not only the forward rate one period before $f_{t-1}^{1,2}$ might have some predictive power, but also forward rates of the preceding periods such as $f_{t-2}^{2,3}$, $f_{t-3}^{3,4}$, etc. The number of lagged forward rates was chosen to be six. Although there is no profound theoretical ground to use exactly this number of lagged forward rates, considering six preceding years should be sufficient for the following analysis. Inclusion of forward rates lying more than six years in the past, although possible, would

[33] A part of this study can be found in Afanasenko/Gischer/Reichling (2011).
[34] This model was initially proposed by Gischer (1997) and Gischer (1998).

probably not be a significant contribution to the empirical results. Therefore, we consider six models each containing an additional lagged forward rate as a predictor of future spot rate:

$$r_t^1 = f\,(f_{t-1}^{1,2}, f_{t-2}^{2,3}, f_{t-3}^{3,4}, f_{t-4}^{4,5}, f_{t-5}^{5,6}, f_{t-6}^{6,7}). \qquad (22)$$

The model according to formula (22) can be represented in the following way:

$$r_t^1 = \beta_0 + \beta_1 f_{t-1}^{1,2} + \beta_2 f_{t-2}^{2,3} + \beta_3 f_{t-3}^{3,4} + \beta_4 f_{t-4}^{4,5} + \beta_5 f_{t-5}^{5.6} + \beta_6 f_{t-6}^{6,7} + \xi_t. \qquad (23)$$

Table 3 provides an overview of the models considered in our study where r_1 and f_i denote the one-year spot rate and a forward rate *i* periods before, respectively.

Table 3: Forward Rate Models

Model	Variables included
1	r_1, f_1
2	r_1, f_1, f_2
3	r_1, f_1, f_2, f_3
4	r_1, f_1, f_2, f_3, f_4
5	r_1, f_1, f_2, f_3, f_4, f_5
6	r_1, f_1, f_2, f_3, f_4, f_5, f_6

3.1.2 Preliminary Data Analysis

The data set employed in this study consists of monthly swap rates for maturities between one and six years over the period 1978 – 2007. The real data on swap rates for maturities from one to six years were available only starting from November 1994. Although the European swap market was fully established in the beginning of the 1980s, the appropriate liquidity for the whole spectrum of maturities was achieved only later. Starting from 1988, the real data on swap rates is available, however, not for all maturities. Thus, swap rates for missing maturities were obtained using linear interpolation. Prior to 1988, no real data on swap rates is available; the required data was estimated through a linear regression approach using yields to maturity of German government bonds. The required spot rates were then computed for maturities from one to six years using equation (5) and recursive computation. Implied forward rates in the six preceding periods were derived from the spot rates.

The period of the financial crisis that occurred in 2008 was not a part of the following analysis for several reasons. At first, including this period could possibly lead to a structural break in the data, which would affect the time series analysis and require a different examination technique. An additional motivation to choose the data set ending in 2007 were significant credit spreads observed in Germany during the financial crisis. Compared to only a few basis points in the pre-crisis period, credit spreads of more than 100 basis points could be observed during the crisis. Such significant credit spreads would bias the empirical results. Finally, as a result of severe liquidity problems, the German swap market experienced a dramatic breakdown, so that the German Federal Bank had to act as the interbank market in this time. Under such circumstances, the assumption of an efficient swap market is not valid for the crisis period. As this study aims at examining a well-functioning fixed-income market, the analysis will be restricted to the pre-crisis period.

Table 4 gives an overview of the basic data characteristics. Already at the first glance it is apparent that forward rates systematically overestimate future spot rates and this effect increases with the lag of forward rates.

Table 4: Descriptive Statistics, Sample 1972 – 2006

Variable	Mean	Median	Max	Min	SD	Skewness	Kurtosis	JB
$\mathbf{r_1}$	5.627	4.950	13.949	2.008	2.546	0.790	3.015	36.316
$\mathbf{f_1}$	6.116	5.552	12.499	2.238	2.308	0.428	2.248	18.870
$\mathbf{f_2}$	6.589	6.352	12.114	2.646	1.979	0.168	2.422	6.500
$\mathbf{f_3}$	6.878	6.875	11.810	2.967	1.793	0.001	2.567	2.730
$\mathbf{f_4}$	7.500	7.507	11.588	4.061	1.616	0.026	2.466	4.184
$\mathbf{f_5}$	7.835	7.801	12.047	4.352	1.526	0.275	2.833	4.816
$\mathbf{f_6}$	8.071	8.034	12.372	4.635	1.422	0.288	3.161	5.199

The difference between lagged forward rates and realized spot rates ranges from 0.5 percent for f_1 to almost 2.5 percent for f_6. Thus, it seems doubtful that forward rates can serve as reliable predictors of future spot rates. Standard deviation of lagged forward rates, however, lies considerably below that of the spot rate and decreases with an increasing lag of forward rates.

Table 4 also presents the results of the Jarque-Berra test, which is a test for normal distribution. A test for normality, Jarque-Berra (JB) statistics, is computed in the following way:[35]

$$\mathrm{JB} = n \cdot \left(\frac{S}{6} + \frac{(K-3)^2}{24} \right), \tag{24}$$

where S represents the third moment of the distribution, the skewness and K corresponds to the fourth moment of the distribution, the kurtosis. The skewness indicates, whether a probability distribution is symmetrical compared to the normal distribution, whose skewness is zero. It can be estimated in the following way:[36]

$$S = \sum_{i=1}^{N} \left(\frac{(R_i - \overline{R})}{\sigma_i} \right)^3 \cdot \frac{1}{N}. \tag{25}$$

The kurtosis characterizes the "thickness" of the tails of a distribution, in comparison with a normal distribution, whose kurtosis is equal to 3. A positive excess kurtosis would mean that the probability of obtaining extreme events is higher than that of a normal distribution, i.e. the distribution would have "fat tails". The kurtosis is given by:[37]

$$K = \sum_{i=1}^{N} \left(\frac{(R_i - \overline{R})}{\sigma_i} \right)^4 \cdot \frac{1}{N}. \tag{26}$$

As kurtosis of a normal distribution equals 3, (K – 3) measures excess kurtosis in formula 24. It can be shown that under the null hypothesis of normal distribution the test statistic has a Chi-square distribution with two degrees of freedom. With the help of the *p*-value of the test statistic it is possible to make a judgment about the normality of the considered distribution. *P*-value shows the probability to obtain the value of the test statistic, which is the same or greater than the observed one. In other words, this value is defined as the lowest significance level at which the null hypothesis can be rejected. Thus, if the *p*-value is sufficiently low, we can reject the null hypothesis.[38]

---

[35] See Jarque/Berra (1987), pp. 163-172.
[36] See Feibel (2003), p. 149.
[37] See Gujarati (1995), pp. 769-771.
[38] See Pindyck/Rubinfeld (1998), pp. 42-43.

As mentioned before, the data set includes real swap rates only starting from 1994. Prior to that year, missing swap rates were obtained with the help of linear interpolation or regression techniques. Because of such a long and mixed data set, we firstly conduct a breakpoint test to identify possible structural breaks. Structural breaks in the data occur because of some structure changes in the relationship among the considered variables. This could be some major regime changes, policy shifts and other events such as wars or natural catastrophes. As a result, the parameters in a regression equation could be different before and after the break. Intercept, slope or both could have changed.[39]

Testing for structural breaks is essential for identifying the time series properties. Testing for unit roots in a data set contaminated by structural breaks can result in misleading conclusions. For example, a unit root test could indicate that the data is non-stationary although it is better characterized as stationary with structural breaks. Perron (1989) was the first to investigate the implications for unit roots in the presence of structural breaks. He states that previous findings that most macroeconomic time series have a unit root are attributable to structural breaks in the data. The unit root tests tend to over-reject the null hypothesis of no unit root. In his study, Perron (1989) could reject the null hypothesis of a unit root in a majority of cases if structural breaks are incorporated.[40]

There is a number of tests to detect structural breaks. One popular test is the Chow breakpoint test;[41] however, it requires a prior specification of the date when the break occurred. The sample is divided into two subsamples, T_1 and T_2 in order to determine, whether they resulted from the same data generating process (DGP). For this purpose, two regressions are run separately and the residual sums of squares are compared with the help of the F-test:

$$\text{Chow test} = \frac{(\bar{u}'\bar{u} - u'u)/T_2}{u'u/(T_1 - H)} \tag{27}$$

where $\bar{u}'\bar{u}$ is the sum of squared residuals from the whole sample regression; $u'u$ represents the sum of squared residuals from the regression using T_1 observations; H is the number of regression coefficients. If the value of the F-statistics exceeds the critical value, the null hy-

[39] See Chatfield (2001), pp. 239-241.
[40] Later, methods were developed specifically for testing for unit roots in the presence of structural breaks. Among these tests are Zivot and Andrews (1992) and Lamsdane and Papell (1997).
[41] See Chow (1960), pp. 591-605.

pothesis that two samples were generated by the same DGP will be rejected. However, it is only possible to apply this test to time series data when the date of the break is known exactly.

A technique that does not require specification of the breakpoint date is the Quandt-Andrews breakpoint test.[42] It is similar to the Chow test; the only difference is that the *F*-statistic is computed between every two points of time throughout the sample period. The point in time where the value of the test is the highest is then selected as the breakpoint date. As reported in table 5, the Quandt-Andrews test indicates multiple breakpoints in the data in 1983, 1992, 1993, and 1995. Although it would be natural to expect the test to identify a break in the period 1991 – 1992, i.e. when the German reunification took place, two more breaks were identified, namely, in 1983 and 1995 in which there were no regime changes or other reforms that could cause such a break. The last break occurred in April 1995 and for this year the value of the test statistics is also the largest. Consequently, a sample starting in May 1995 and ending in October 2006 was chosen to analyze the predictive ability of forward rates. This subset is of special interest for us, as it is free of structural breaks. An additional motivation for this choice is the fact that the sample 1995 – 2006 is composed of real data on swap rates. The data for the last 12 months will not be considered in our analysis in order to evaluate an out-of-sample performance of the model. Series plots are shown in figure 7.

Table 5: Quandt-Andrews Breakpoint Test, Sample 1978 – 2006

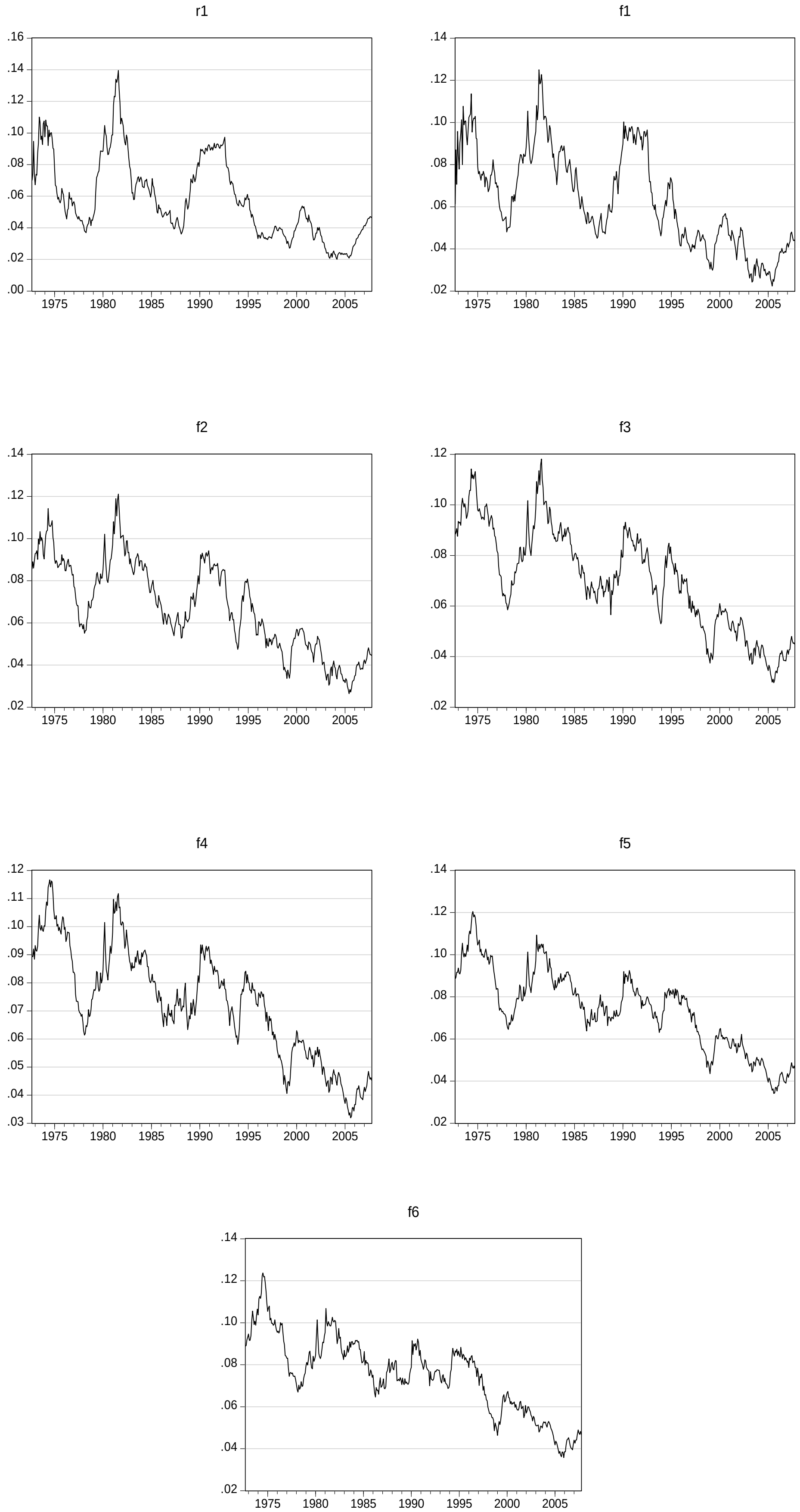
(H_0: No structural breaks within data)

Model	Critical Value	Test Value	Date
r_1, f_1	17.5	45.31*	1983M03
r_1, f_1, f_2	28.6	65.08*	1995M04
r_1, f_1, f_2, f_3	22.7	51.29*	1993M02
r_1, f_1, f_2, f_3, f_4	22.1	48.77*	1983M07
r_1, f_1, f_2, f_3, f_4, f_5	18.5	43.61*	1992M10
r_1, f_1, f_2, f_3, f_4, f_5, f_6	15.2	36.89*	1992M10

*Indicates rejection of the hypothesis at the one percent significance level

[42] The test was introduced by Quandt (1960) and later further developed by Andrews (1993).

Figure 7: Series Plot – One-Year Spot Rate and Lagged Forward Rates



The issue of stationarity has received great attention in the past two decades. A stationary process is characterized by mean, variance, and autocovariance which are time-independent. Formally, time series Y_t is (weakly) stationary if the following conditions hold:[43]

$$\begin{aligned} &\mathrm{E}(Y_t) = z_Y && \text{for all } t \\ &\mathrm{Var}(Y_t) = \sigma_Y^2 && \text{for all } t \\ &\mathrm{Cov}(Y_t, Y_{t+k}) = \gamma_k && \text{for all } t \text{ and } k \end{aligned}$$

where μ_Y, σ_Y^2 and γ_k are constants. In general, ordinary least squares (OLS) estimation is only justified when the data exhibits constant mean and variance. Non-stationary time series or, put differently, time series having a unit root are characterized by means and variances that are time-dependent. Thus, false inferences from conventional statistics could be drawn when OLS techniques are employed.[44] Estimating non-stationary time series with OLS could lead to meaningless results or "spurious" regression, in which relationships are confirmed among variables completely unrelated to each other.[45] Conventional tests statistics cannot be applied as their asymptotic distributions are non-standard under non-stationarity.

There is an ongoing discussion on the time series properties of interest rates. Using the Dickey-Fuller[46] (DF) test has been standard practice to test for the presence of a unit root in the empirical literature, which generally resulted in the inability to reject the null hypothesis that interest rates are non-stationary time series. Suppose that the time series is represented by a first-order autoregressive process, or AR(1):

$$r_t = z + \rho r_{t-1} + \varepsilon_{t,} \qquad \varepsilon \sim \mathrm{N}(0, \sigma_\varepsilon^2) \tag{28}$$

The DF test, which aims at checking whether $\rho < 1$, proceeds in the following way:

$$\Delta r_t = z + (\rho - 1) r_{t-1} + \varepsilon_t = z + \vartheta r_{t-1} + \varepsilon_t, \tag{29}$$

[43] See Thomas (1997), pp. 373-375.
[44] See Asteriou/Hall (2006), pp. 291-292.
[45] Granger and Newbold (1974) and Granger and Newbold (1977) run a number of regressions that involved non-stationary time series that were completely unrelated to each other. In 75 percent of all cases, they found evidence of statistically significant relationships.
[46] See Dickey and Fuller (1979).

where $\vartheta = \rho - 1$. The null hypothesis of the DF test is $\vartheta = 0$, i.e. the time series has a unit root. Equation (29) is estimated using OLS and the obtained *t*-values are compared with the critical values reported by Dickey and Fuller (1979). Of course, the autoregressive process can also be of a higher order. In this case, the DF test is augmented by additional lags and is referred to as the augmented Dickey-Fuller (ADF) test. It also has three different specifications, depending on inclusion or non-inclusion of a time trend and constant. Equation (30) gives an example of the ADF test including a constant term; equation (31) shows the ADF test with both constant and trend; finally, equation (32) represents the least restrictive version of the test, with neither a constant nor a trend term:[47]

$$\Delta r_t = z + \vartheta r_{t-1} + \sum_{i=1}^{\rho} \varphi_i \Delta r_{t-1} + \varepsilon_t \tag{30}$$

$$\Delta r_t = z + \omega T + \vartheta r_{t-1} + \sum_{i=1}^{\rho} \varphi_i \Delta r_{t-1} + \varepsilon_t \tag{31}$$

$$\Delta r_t = \vartheta r_{t-1} + \sum_{i=1}^{\rho} \varphi_i \Delta r_{t-i} + \varepsilon_t . \tag{32}$$

The number of lags in the above equations has to be selected appropriately in order to capture the nature of the process, but at the same time also not to include too many lags as this would lead to the loss of degrees of freedom. Commonly, the Akaike or the Schwarz information criterion is used to select the number of lags. The ADF test faced criticism because of having too low power to reject the null hypothesis of a unit root.[48] As a response, modifications of the ADF tests were elaborated, including the Philips-Perron (PP) test[49] and the Dickey-Fuller generalized least squares (GLS) test.[50] If, according to the above stationarity tests, a time series has to be differenced *d* times in order to become stationary, it is said to be integrated of order *d*, denoted *I(d)*. If a series turns out to be stationary and does not require differencing, it is referred to as integrated of order zero *I*(0).

It is common to consider interest rate data as well as many other macroeconomic data, to be an *I(1)* process, i.e. data become stationary after first-differencing. This view became especially popular after the study of Nelson and Plosser (1982), who conducted unit root tests on a

[47] See Patterson (2010), pp. 218-223.
[48] See DeJong et al. (1992).
[49] See Philips and Perron (1988).
[50] See Elliot/Rothenberg/Stock (1996), pp. 813-836.

number of US economic time series and found almost all series to be non-stationary. Later, Zhang (1993) examined US interest rates for maturities from one month to 30 years and confirmed the presence of a unit root for all considered maturities. The study of Tabak (2009) found Brazilian interest rates of maturities below one year to be non-stationary time series.[51]

However, some authors question non-stationarity of interest rates. For example, Wu and Chen (2000) mention that the low power of standard techniques, such as ADF or PP tests in small samples could be the reason for the inability to reject the stationarity hypothesis in interest rate data. The authors use the DF GLS test and report the rejection of the unit root hypothesis for two out of seven countries. Beechey (2009) criticizes the common acceptance of non-stationarity of interest rates. He stresses the inability of standard testing methods to distinguish between a pure unit root and near unit root process when the root is close, but not equal to unity. [52] A shortcoming of the ADF and PP as well as DF GLS tests is that they are based on a unit root assumption and thus, unless there is very strong evidence against the null hypothesis, it tends to be accepted. A new procedure for testing for stationarity was suggested by Kwiatkowski, Philips, Schmidt and Shean (1992). The unit root test, which is known as the KPSS test, is considered to be more powerful as its null hypothesis is a stationary process instead of a unit root process. This test is, therefore, less likely to reject stationarity. In order to obtain as accurate results as possible concerning the time series properties of our data, we employ several unit root tests including those whose null hypothesis is the absence of a unit root. The results on the ADF test with different specifications are displayed in table 6; table 7 contains the outcomes of the PP, DF GLS and KPSS unit root tests.

As table 6 and 7 show, unit root tests deliver quite uniform results in the considered sample. If variables are expressed in levels, the null hypothesis of a unit root process fails to be rejected by all tests for all time series. This is also true for all three specifications of the ADF test, whose results are not affected by the inclusion of a constant or a time trend. Most interesting is a comparison between the traditional unit root tests with the null hypothesis of a unit root from one side and the KPSS test with the null hypothesis of a stationary process from the other side. However, we observe that even this more powerful test rejects the null hypothesis of a stationary process at least at the five percent significance level. In case of f_1, f_2, f_3, f_4, and f_5

[51] For further findings in favor of interest rates containing a unit root see, among others, Campbell and Shiller (1987), Hall et al. (1992), Shea (1992).

[52] Lanne (2000) and Beechey et al. (2009) suggest an alternative technique to test for stationarity in case of near-integrated processes.

the null hypothesis is rejected even at the one percent significance level. Therefore, we find evidence that interest rates are in fact the best described by a unit root process.

Regarding the order of integration of the data, when interest rates are expressed in first differences, ADF and PP test both reject the null hypothesis of a unit root process at the one percent significance level. The DF GLS test supports this result at the one percent significance level for all cases excluding f_2 and f_6, where rejection is at the five percent significance level. The KPSS test reinforces the conclusion obtained from the previous tests. It fails to reject the null hypothesis of a stationary process.

Table 6: ADF Unit Root Tests on Forward Rates, Sample 1995 – 2006

	Level			First difference		
	ADF with constant and trend	ADF with constant	ADF without constant and trend	ADF with constant and trend	ADF with constant	ADF without constant and trend
r_1	–1.98	–2.22	–0.85	–5.88*	–5.79*	–5.81*
f_1	–2.58	–2.05	–1.33	–9.82*	–9.84*	–9.82*
f_2	–2.15	–1.33	–1.19	–9.69*	–9.72*	–9.71*
f_3	–2.25	–1.12	–1.07	–10.74*	–10.78*	–10.76*
f_4	–2.04	–1.48	–1.48	–11.24*	–11.28*	–11.21*
f_5	–1.94	–1.56	–1.45	–12.32*	–12.35*	–12.26*
f_6	–2.40	–1.36	–0.46	–14.13*	–14.12*	–14.16*

Notes: *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively. The null hypothesis of the ADF test is a unit root process.

Table 7: PP, DF GLS and KPSS Unit Root Tests, Sample 1995 – 2006

	Level			First difference		
	PP	DF GLS	KPSS	PP	DF GLS	KPSS
r_1	–2.09	–1.33	0.540**	–9.42*	–3.93*	0.142
f_1	–1.87	–0.77	0.884*	–9.87*	–3.54*	0.048
f_2	–1.77	–0.47	0.919*	–9.82*	–2.48**	0.037
f_3	–1.40	–0.57	1.015*	–10.89*	–4.05*	0.050
f_4	–1.60	–0.06	1.083*	–5.88*	–11.33*	0.060
f_5	–1.60	–0.10	1.050*	–12.33*	–11.49*	0.082
f_6	–1.39	–1.45	0.730**	–13.86	–2.47**	0.131

Notes: *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively. The null hypothesis of the KPSS test is a stationary process. All other tests assume a unit root process. The information criterion used in DF tests is that of Schwarz, as the usage of the Akaike criterion resulted in a higher number of lags. KPSS test with Bartel kernel was applied.

It was also of interest to study the time-series properties of the whole sample, for the period 1972 – 2007. Based on the ADF test, table 8 once more confirms that both spot and forward rates are non-stationary time series in levels and stationary in first differences. The only exception constitutes f_1, for which the ADF test with constant and trend rejects the null hypothesis of a unit root process. However, this result holds only at the ten percent significance level. This finding is reinforced in table 9, where the results of the PP, DF GLS and KPSS test stress the non-stationary character of the considered time series. Only DF GLS test rejects the hypothesis of unit root process for r_1 and f_1 at the ten percent and five percent significance level, respectively. For f_1, the DF GLS test rejects the null hypothesis of a unit root process for the data in levels. However, when first differences are taken, the test fails to reject the null, indicating a process that is stationary in levels, but non-stationary in first differences. Thus, all time series will be treated as a first-difference stationary process.

Table 8: ADF Unit Root Tests, Sample 1972 – 2006

	Level			First difference		
	ADF with constant and trend	ADF with constant	ADF without constant and trend	ADF with constant and trend	ADF with constant	ADF without constant and trend
r_1	–2.41	–2.21	–1.38	–18.17*	–18.19*	–18.20*
f_1	–3.32***	–1.99	–0.84	–24.75*	–24.79*	–24.81*
f_2	–3.03	–1.73	–1.09	–18.41*	–18.43*	–18.44*
f_3	–2.95	–1.42	–1.05	–19.29*	–19.31*	–19.31*
f_4	–2.88	–1.27	–1.06	–19.29*	–19.31*	–19.30*
f_5	–3.03	–1.25	–1.03	–20.49*	–20.50*	–20.50*
f_6	–3.13	–1.32	–1.01	–21.11*	–21.13*	–21.13*

Notes: *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively. The null hypothesis of the ADF test is a unit root process.

Table 9: PP, DF GLS and KPSS Tests, Sample 1972 – 2006

	Level			First Difference		
	PP	DF GLS	KPSS	PP	DF GLS	KPSS
r_1	–2.05	–1.84***	1.14*	–18.46*	–1.79***	0.046
f_1	–1.94	–2.00**	1.54*	–24.44*	–0.29	0.063
f_2	–1.66	–1.04	1.84*	–18.44*	–2.78*	0.037
f_3	–1.52	–0.72	1.93*	–19.29*	–18.51*	0.037
f_4	–1.40	–0.58	1.92*	–19.31*	–19.32*	0.041
f_5	–1.32	–0.64	1.92*	–20.51*	–19.08*	0.047
f_6	–1.36	–0.71	1.89*	–21.12*	–20.99*	0.047

Notes: *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively. The null hypothesis of the KPSS test is a stationary process. All other tests assume a unit root process. The information criterion used in DF tests is that of Schwarz, as the usage of the Akaike criterion resulted in a higher number of lags. KPSS test with Bartel kernel was applied.

3.1.3 Cointegration and the Error-Correction Model

The ordinary least squares technique is not appropriate in the case of integrated time series. One possible approach in the case of an $I(1)$ process is to use first-differenced series. Suppose we want to estimate the following equation:

$$r_t = \beta_0 + \beta_1 f_{t-1} + \phi_t . \tag{33}$$

If r_t and f_t are $I(1)$ series, then one can estimate:

$$\Delta r_t = a_0 + a_1 \Delta f_t + \phi_t \tag{34}$$

to achieve stationarity. However, this procedure is not very popular; critics of this approach point out that the "long-run" information could be ignored in the case that the series are expressed in first differences, and not in levels.[53] A possible solution to this problem would be to incorporate some kind of long-run information into the above formula and estimate the following expression:

[53] For further details, see Davidson et al. (1978).

$$\Delta r_t = \delta_0 + \delta_1 \left(r_{t-1} - \hat{\beta}_1 f_{t-1} - \hat{\beta}_0 \right) + a_1 \Delta f_t + \phi_t . \tag{35}$$

Equation (35) incorporates long-term as well as short-term parameters and is referred to as an error correction model (ECM).[54] The term in brackets, $r_{t-1} - \hat{\beta}_1 f_{t-1} - \hat{\beta}_0$, is an error correction term and the coefficient δ_1 measures the speed of adjustment to correct this error. A short-coming of this approach is that, although two differenced terms are stationary, the error correction term is a linear combination of non-stationary variables and is, therefore, also non-stationary.[55]

One way to address this problem is to apply the concept of cointegration which was introduced by Granger (1987). If two time series are $I(1)$ but a linear combination of them can be found that is stationary, i.e. $I(0)$, then these series are said to be cointegrated. This result can also be extended for the case of more than two series and a higher order of integration. If X_t is a vector of $I(d)$ variables and a vector $\tilde{\alpha} \neq 0$ exists so that linear combination $\tilde{\alpha}' X_t \sim I(d-b)$, $b > 0$, then the components of the vector X_t are cointegrated of order d,b denoted $CI(d,b)$.[56] Vector $\tilde{\alpha}$ is known as the cointegrating vector (CIV), $\tilde{\alpha}' X_t$ is a vector of error correction terms.

Cointegration represents a useful tool for analyzing the data, as it allows for capturing possible long-run relationships between non-stationary time series. If cointegration is present, time-series, although non-stationary, are linked to each other through a common trend. Stock (1987) demonstrated that for cointegrated variables, the OLS estimator of the CIV, $\hat{\tilde{\alpha}}$, will be "super consistent", i.e. it will converge to the true parameter value at a faster rate than the OLS estimator of a regression involving stationary variables. However, its distribution will be non-standard and therefore conventional statistical inference is not applicable. One crucial implication of cointegration, known as the Granger representation theorem,[57] is that if the series are $CI(1,1)$, an ECM will be a valid representation of the data. The error correction representation is appealing because it only contains stationary variables, as the term in brackets

[54] The concept of an ECM was first mentioned by Sargan (1964). Later, it was further developed by Hendry and Anderson (1977).
[55] See Holden/Perman (2007), pp. 64-66.
[56] Flores and Szafarz (1996) provide an enlarged definition of cointegration and show that cointegration may also arise among the series with different order of integration.
[57] See Granger (1983), Engle and Granger (1987).

in equation (35) will be stationary under cointegration. Currently there are several tests available to detect cointegration. In the case that only two time series are involved, the Engle-Granger two-step estimation procedure[58] can be applied. As a first step, equation (33) is estimated using OLS and the error term is calculated:

$$e_t = r_t - \hat{\beta}_0 - \hat{\beta}_1 f_{t-1}. \tag{36}$$

Subsequently, the error term from the regression is tested for stationarity using one of the standard tests:[59]

$$e_t = \rho e_{t-1} + \upsilon_t. \tag{37}$$

If the error term is stationary, the series are said to be cointegrated. Applied to a multivariate case, the Engle-Granger two-step procedure can no longer guarantee the uniqueness of the estimated CIV as there could exist $p-1$ linear relationships in the case of p involved series.[60]

Jonansen (1988, 1991) and Johansen and Juselius (1990) developed a technique that enables detection and estimation of multiple CIVs.[61] Being a part of econometric software packages, this procedure is frequently applied to test for cointegration. At first, the system is represented as a vector autoregression (VAR) model of order k:[62]

$$X_t = \Pi_1 X_{t-1} + ... + \Pi_k X_{t-k} + \varepsilon_t, \tag{38}$$

where X_t is a p-dimensional vector of $I(1)$ series, Π_i, i=1….k is a $p \times p$ matrix of coefficients and ε_t is a $p \times 1$ vector of error terms that are independently identically distributed (i.i.d.). The VAR models are models without strong theoretical basis that do not distinguish between exogenous and endogenous variables. The above expression can also be represented in the error correction form:[63]

58 See Engle/Granger (1987).
59 Such as DF, DF GLS, PP or KPSS test.
60 See Harris/Sollis (2005), pp. 92-93.
61 Another multivariate cointegration test was proposed by Stock and Watson (1988).
62 See Lütkepohl/Krätzig (2004), pp. 88-89.
63 See Asteriou/Hall (2006), pp. 319-320.

$$\Delta X_t = \Gamma_i \Delta X_{t-1} + \ldots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + \varepsilon_t, \tag{39}$$

where

$$\Gamma_i = -I + \Pi_1 + \ldots + \Pi_i$$
$$\Pi = -I + \Pi_1 + \ldots + \Pi_k \tag{40}$$

and I denotes the identity matrix. In equation (39), Π is the only term that is expressed in levels. Π equal to zero indicates the absence of cointegration; Π having a full rank implies that the involved series are all stationary. Therefore, tests for cointegration focus on determining whether matrix $\Pi = \tilde{\beta}\tilde{\alpha}'$ has a reduced rank, which can be at most equal to p–1. $\tilde{\alpha}'$ denotes the $p \times r$ matrix of CIVs whereas $\tilde{\beta}$ is the $p \times r$ matrix of adjustment coefficients. The maximum likelihood estimator, which was developed by Johansen (1988), is applied to calculate the eigenvalues of Π.[64]

Two types of tests are carried out to test for the number of cointegrating vectors or, put differently, for the number of the long-run relationship among the series. The maximum eigenvalue test, which is used to test the null hypothesis of r CIVs against the alternative of that their number is r+1, is based on the following test statistics:

$$\lambda_{\max} = -T \ln\left(1 - \lambda_{r+1}\right), \tag{41}$$

where T is the number of observations and λ_r is an eigenvalue associated with the cointegrating vector r. The second cointegration test, the trace test, is based on the following:

$$\lambda_{trace} = -T \sum_{i=r+1}^{k} \ln\left(1 - \lambda_i\right). \tag{42}$$

The null hypothesis of the trace test is that there are at most r CIVs with the alternative hypothesis being the existence of more than r CIVs. Critical values obtained by Johansen and

[64] See Asteriou/Hall (2006), pp. 321.

Juselius (1990) should be applied to assess the results of these tests. If in a system of *p* series *r* cointegrating vectors are found, the system is said to be driven by *p*–*r* common trends.[65]

In the contest of the EH of the term structure, cointegrated interest rates of different maturities are usually interpreted as an evidence in favor of the EH. A large number of CIVs in the system is also an indicator of the validity of the EH. However, for a pure EH to be valid, it has to be tested, whether the CIV is different from unity.

Due to its theoretical appeal, the cointegration framework has been applied in the term structure literature. Some early studies in the beginning of the 1990s have focused on determining whether US term structure components are cointegrated. Hall, Anderson and Granger (1992) use T-bill data for the period 1970 – 1988 with one to eleven months maturity to test for cointegration among all eleven yields as well as pairwise between different yields. According to the authors, for the EH to hold there should be *p*–1 CIVs for a set of *p* series. They find that eleven interest rates have ten CIVs, i.e. the EH holds. In contrast, Zhang (1993) examines 19 interest rate series dating from 1964 to 1986 and documents the existence of 16 cointegrating vectors, or three common trends, in the US term structure.

The findings of Engsted and Tanggaard (1994), who analyze the US term structure using a data set ranging 1952 – 1987, also indicate one common trend in the data. Dominguez and Novales (2002) apply cointegration analysis to estimate interest rates with one, three and six month maturity depending on one lagged forward rate. Their sample includes US, British, Japanese, Spanish, French, Italian, Swiss and German data for the period 1978 – 1998. With several exceptions, they found cointegration between the pairs of interest rates and conclude that forward rates can serve as unbiased predictors of future spot rates. In addition, they find that lagged forward rates can predict future spot rates better than the univariate autoregressive model.

Ghazali and Low (2002) investigate the case of Malaysia and show that long-and short-term interest rates are cointegrated. The ECM elaborated by the authors has significant adjustment coefficients in most cases. They analyze both long-short and short-long models and find that, based on the absolute size of the adjustment coefficients, long-term rates have stronger power

[65] See Burke/Hunter (2005), pp. 89-90.

in determining the short-term rates. Finally, Tabak (2009) uses Brazilian swap rates for one, three, six and 12 months ranging from 1995 to 2006 to test the expectations hypothesis using term spreads. Although cointegration is present, the study rejects the pure form of the EH. Evidence for European countries using cointegration analysis is quite scarce. We intend to close this gap providing the cointegration analysis for the German term structure.

3.2 Empirical results

This section contains the results of the cointegration and the error correction analysis for the considered models. This analysis is followed by forecasts produced by the models for the period from November 2006 to October 2007. Finally, forecasting performance is evaluated with the help of mean absolute error (MAE), root mean square error (RMSE) and Theil's inequality coefficient.

3.2.1 Cointegration analysis

For the first model, which involves only the one-year spot rate and one lagged forward rate, we employ the Engle-Granger two-step procedure as well as the Johansen approach. However, for the remaining models under consideration we employ Johansen cointegration tests, as this procedure allows identifying all relevant CIVs. An important issue when using Johansen cointegration tests is the choice of the number of lags in VAR. If this number is too low, the model is specified incorrectly. From the other side, if there are too many lags, this leads to the loss of degrees of freedom.[66]

The optimal number of lags can be selected on the basis of various information criteria, such as Akaike Information Criterion (AIC)[67], Schwarz Information Criterion (SIC)[68], Final Prediction Error (FPE), sequential modified likelihood ratio (LR) test statistic, and Hannan-Quinn criterion (HQ).[69] The number of lags indicated by each test is reported in table 10 for all considered models. It shows that AIC, SIC, FPE and HQ criterion produce uniform results regarding the number of lags in VAR, which is especially evident for model 2 and 4 to 6. In contrast, the LR test statistic indicates significantly larger number of lags. Thus, cointegration

[66] See See Asteriou/Hall (2006), p. 322.
[67] See Akaike (1974).
[68] See Schwarz (1979).
[69] See Hannan/Quinn (1979).

tests are performed several times, each time changing the number of lags, in accordance with the respective lag order selection criterion.[70]

Table 10: Lag Order Selection Criteria[71]

Model	LR	AIC	SC	HQ	FPE
1	7	2	1	1	2
2	8	1	1	1	1
3	7	2	1	1	2
4	7	1	1	1	1
5	7	1	1	1	1
6	7	1	1	1	1

Table 11 and 12 report the results of cointegration tests for the six models with the number of lags, *L*, selected by AIC and LR, respectively.[72] For the models considered in table 11 both tests reject the null hypothesis of no cointegration in case of models 2, 4, and 5 at the five percent significance level. For model 1 the null hypothesis is rejected even at the one percent significance level. However, the results are not so straightforward for model 3 and 6. Whereas the maximum eigenvalue test rejects the null hypothesis of no cointegration at the five percent level, the trace statistics is not able to verify this result. The situation is reversed for model 6. The trace test indicates cointegration at the five percent level while the maximum eigenvalue test is (just) not able to reject the null. For all models the number of cointegrating vectors is found to be one. However, if the number of lags selected by the LR test were applied, as shown in table 12, a higher number of CIVs was detected by the tests. This number ranges from one for the first model and reaching four for model 6.

[70] However, the results are insensitive to inclusion of more lags, i.e. cointegration is still found at the 5percent level. Thus, we consider Johansen's tests to be an appropriate procedure in our case.

[71] The models indicated in the table correspond to those introduced on page 33. Model 1 includes one lagged forward rate; each of the subsequent models includes one additional lagged forward rate.

[72] All tests were conducted under no deterministic trend assumption. It is worth mentioning that the results of cointegration tests are not affected by introducing the trend assumption.

Table 11: Johansen Cointegration Tests, Sample 1995 – 2006

Model	Hypothesis	λ_{trace}	Critical value	L	$\lambda_{\max}$	Critical value
1	r=0	31.56*	25.08	2	24.48*	20.16
	$r \geq 1$	7.08	12.76		7.08	12.76
2	r=0	48.61**	35.19	1	31.34**	22.30
	$r \geq 1$	17.27	20.26		12.98	15.89
3	r=0	50.48	54.08	2	29.00**	28.58
	$r \geq 1$	21.48	35.19		12.08	22.30
4	r=0	79.78**	76.97	1	38.66**	34.80
	$r \geq 1$	41.12	54.08		19.15	28.59
5	r=0	106.94**	103.84	1	45.69**	40.96
	$r \geq 1$	61.24	76.97		21.18	34.81
6	r=0	144.81**	134.68	1	47.07	47.08
	$r \geq 1$	97.74	103.85		37.00	40.96

Notes: *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively; L denotes the number of lags in VAR in levels.

In case of model 2 to 4, the number of CIVs detected by both tests is two. It is also worth noting that trace and maximum eigenvalue test indicate cointegration for all models and yield an equal number of CIVs in all but one case. The only exception is model 5, where trace test indicates that three CIVs are present while the maximum eigenvalue test detects only two of them. Thus, including a larger number of lags into the VAR model does not affect the main result: in all models variables remain cointegrated. The difference is that more cointegrating vectors are detected. A similar situation is observed when the full data set available for the period 1978 – 2006 is analyzed. Cointegration is also found in this sample; however, the number of CIVs is more than just one for several models, which is reflected in table 13.

Table 12: Johansen Cointegration Tests, Sample 1995 – 2006

Model	Hypothesis	λ_{trace}	Critical value	L	$\lambda_{\max}$	Critical value
1	r=0	35.81*	19.94	7	31.91*	18.52
	$r \geq 1$	3.89	6.63		3.89	6.63
2	r=0	92.99*	35.46	8	55.43*	25.86
	$r \geq 1$	37.56*	19.94		35.53*	18.5
	$r \geq 2$	2.04	6.63		2.04	6.63
3	r=0	76.38*	54.68	7	39.97*	32.72
	$r \geq 1$	36.41*	35.46		26.85*	25.86
	$r \geq 2$	9.55	19.94		8.48	18.52
4	r=0	100.95*	69.81	7	44.12*	33.88
	$r \geq 1$	56.83*	47.86		30.07*	27.58
	$r \geq 2$	26.76	29.80		16.53	21.13
5	r=0	148.20*	104.96	7	52.36*	45.87
	$r \geq 1$	95.85*	77.82		38.91	39.37
	$r \geq 2$	56.94*	54.68			
	$r \geq 3$	29.01	35.46			
6	r=0	211.31*	135.97	7	54.79*	52.31
	$r \geq 1$	156.51*	104.96		54.41*	45.87
	$r \geq 2$	102.10*	77.82		44.96*	39.37
	$r \geq 3$	57.14*	54.68		28.69	32.72

Notes: *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively; L denotes the number of lags in VAR in levels.

Table 13: Johansen Cointegration Tests, Sample 1978 – 2006

Model	Hypothesis	λ_{trace}	Critical value	L	$\lambda_{\max}$	Critical value
1	r=0	39.36*	19.94	1	35.22*	18.52
	$r \geq 1$	4.14	6.63		4.14	6.63
2	r=0	75.09*	35.46	1	39.81*	25.86
	$r \geq 1$	35.26*	19.94		32.91*	18.5
	$r \geq 2$	2.35	6.63		2.35	6.63
3	r=0	98.70*	54.68	1	40.93*	32.72
	$r \geq 1$	57.76*	35.46		35.68*	25.86
	$r \geq 2$	22.08*	19.94		20.08	18.52
	$r \geq 3$	1.99	6.63		1.99	6.63
4	r=0	197.03*	77.82	7	89.81*	39.37
	$r \geq 1$	107.22*	54.68		75.19*	32.72
	$r \geq 2$	32.04	35.46		21.78	25.86
5	r=0	269.12*	104.96	7	89.01*	45.87
	$r \geq 1$	180.11*	77.82		79.38*	39.37
	$r \geq 2$	100.74*	54.68		69.57*	32.72
	$r \geq 3$	31.17	35.46		20.52	25.86
6	r=0	211.31*	135.97	7	54.79*	52.31
	$r \geq 1$	156.51*	104.96		54.41*	45.87
	$r \geq 2$	102.10*	77.82		44.96*	39.37
	$r \geq 3$	57.14*	54.68		28.69	32.72

Notes: *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively; L denotes the number of lags in VAR in levels.

It was also of interest to test for cointegration in pairs, i.e. including the spot rate and each lagged forward rate. The corresponding results are shown in table 14.

Table 14: Johansen Cointegration Tests between Pairs of Spots and Forward Rates, Sample 1995 – 2006

Variables	Hypothesis	λ_{trace}	Critical value	L	$\lambda_{\max}$	Critical value
r_1, f_1	r=0	31.79*	25.08[a]	2	27.64*	20.16[a]
r_1, f_2	r=0	9.07	17.98[b]	4	5.31	13.91[b]
r_1, f_3	r=0	8.11	17.98[b]	4	5.32	13.91[b]
r_1, f_4	r=0	8.06	17.98[b]	2	5.90	13.91[b]
r_1, f_5	r=0	8.40	17.98[b]	2	5.52	13.91[b]
r_1, f_6	r=0	17.56	17.98[b]	3	13.85	13.91[b]

Notes: * denotes the rejection of the null hypothesis at the 1% level; L denotes the number of lags in VAR in levels; a and b are 1% and 10% critical values, respectively.

From table 14 one can observe that cointegration for the pairs of spot and lagged forward rates only holds in the first case, for r_1 and f_1.[73] For all other pairs both cointegration tests are not able to reject the hypothesis that variables are not cointegrated even at the ten percent level. There is some ambiguity for the last case, r_1 and f_6, where both tests are (just) not able to reject the null hypothesis.

Thus, we have found that there is a long-run relationship among spot rates and lagged forward rates. In this case it could be possible to use this information for constructing forecasts. Table 15 shows the normalized cointegrating coefficients for the sample 1995 – 2006. The results of the maximum likelihood estimation of parameters, as opposed to the conclusions drawn from the cointegration tests, are not so promising. One can infer from table 15 that only f_1, i.e. the forward rate determined one period before, is highly significant in all six models and, therefore, exhibits explanatory power with respect to the future one-year spot rate. However, it has a positive sign only for models 1 through 4. In models 5 and 6 the wrong direction is predicted by the lagged forward rate f_1. Forward rates which prevailed two and three periods before, f_2 and f_3, are either insignificant or significant only at the ten percent level. In most of the considered models they cannot contribute to the prediction of the future spot rate. Up to model 5,

[73] When the Engle-Granger two-step procedure is applied to check if r_1 and f_1 are cointegrated, the KPSS test is not able to reject the null hypothesis of a stationary process. The DF-GLS test confirms this result, rejecting the null of a unit root process at the five percent level. In contrast, the outcome of the ADF and the PP test is a non-stationary error term. As the Johansen approach indicates cointegration at the one percent level, these results may be interpreted as an evidence of poor performance of the ADF and the PP test.

f_2 also has a negative sign. Although f_3 has a positive sign in models 3 and 4, it becomes negative starting from model 5.

Table 15: Parameter Estimates, Sample 1995 – 2006

Model	Parameter estimates						
	Constant	β					
1	–0.015 (0.00979) [–1.55]	1.2316* (0.2241) [5.49]					
2	–0.0109 (0.01393) [0.78]	1.662* (0.261) [6.36]	–0.476*** (0.265) [–1.79]				
3	–0.007 (0.014) [–0.05]	1.645* (0.268) [6.12]	–0.554 (1.004) [–0.53]	0.0134 (1.008) [–0.013]			
4	–0.0009 (0.0027) [–0.034]	3.050* (0.459) [6.64]	–3.353*** (1.829) [-1.83]	2.95*** (1.817) [1.62]	–1.338** (0.486) [2.76]		
5	–0.0173 (0.031) [–0.55]	–3.99* (0.558) [–7.16]	2.25 (1.978) [1.13]	–2.20 (1.958) [-1.12]	1.204*** (0.607) [1.98]	2.099* (0.65) [3.22]	
6	–0.0211 (0.012) [–1.79]	–1.441* (0.186) [–7.76]	–0.024 (0.633) [–0.04	–0.189 (0.618) [–0.30]	0.393** (0.194 [2.03]	0.84* (0.217) [3.686]	0.579* (0.203) [2.85]

Notes: numbers in parentheses stand for the maximum likelihood standard errors; numbers in square brackets are *t*-statistics; *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively

Whereas f_2 and f_3 seem to have no forecasting ability with respect to the future spot rate, forward rates lying farther in the past might be more useful in explaining the spot rate. In models 4, 5 and 6, f_4 is significant at the five or ten percent level. In models 5 and 6, f_5 is significant at the one percent level whereas f_6 also seems to have some predictive power in model 6 at the one percent level. Both f_5 and f_6 have positive signs. To summarize, for models 4, 5 and 6 the first and the last forward rates are significantly different from zero at least at the five percent significance level. Thus, we found evidence that forward rates one period before as well as forward rates that are lying five and six periods before contain some explanatory power regarding the one-year spot rate, although the sign reversion starting from model 5 is puzzling. Although sign reversion starting from model 5 could be interpreted as mean reversion according to business cycles reflected in the level of interest rate, such interest rate cycles show no constant length. Forward rates lying in the "middle" f_2, f_3 and f_4 do not seem to be a useful tool in forecasting spot rates. While cointegration analysis represents a more sophisticated technique to analyze time series data, it appears to be interesting to compare its results with a simple model which treats time series as first differences. This way, although not very popular, as cointegration analysis is strongly preferred in the recent literature, represents an easy way to deal with time series data. Table 16 demonstrates the results of the model 1 to 6 with forward rates expressed as first-differences:

$$\begin{aligned}
\Delta r_t^1 &= \beta_0 + \beta_1 \Delta f_{t-1}^{1,2} + \xi_t \\
\Delta r_t^1 &= \beta_0 + \beta_1 \Delta f_{t-1}^{1,2} + \beta_2 \Delta f_{t-2}^{2,3} + \xi_t \\
\Delta r_t^1 &= \beta_0 + \beta_1 \Delta f_{t-1}^{1,2} + \beta_2 \Delta f_{t-2}^{2,3} + \beta_3 \Delta f_{t-3}^{3,4} + \xi_t \\
\Delta r_t^1 &= \beta_0 + \beta_1 \Delta f_{t-1}^{1,2} + \beta_2 \Delta f_{t-2}^{2,3} + \beta_3 \Delta f_{t-3}^{3,4} + \beta_4 \Delta f_{t-4}^{4,5} + \xi_t \\
\Delta r_t^1 &= \beta_0 + \beta_1 \Delta f_{t-1}^{1,2} + \beta_2 \Delta f_{t-2}^{2,3} + \beta_3 \Delta f_{t-3}^{3,4} + \beta_4 \Delta f_{t-4}^{4,5} + \beta_5 \Delta f_{t-5}^{5,6} + \xi_t \\
\Delta r_t^1 &= \beta_0 + \beta_1 \Delta f_{t-1}^{1,2} + \beta_2 \Delta f_{t-2}^{2,3} + \beta_3 \Delta f_{t-3}^{3,4} + \beta_4 \Delta f_{t-4}^{4,5} + \beta_5 \Delta f_{t-5}^{5,6} + \beta_6 \Delta f_{t-6}^{6,7} + \xi_t
\end{aligned} \tag{43}$$

Table 16: Parameter Estimates for the Models in First Differences, Sample 1995 – 2006

Model	Parameter estimates						
	β_i						R^2(adj.)
1	–0.052 (0.054) [–0.96]						0.0058
2	–0.052 (0.056) [–1.05]	–0.028 (0.054) [–0.52]					0.00039
3	–0.065 (0.056) [–1.16]	–0.118 (0.125) [–0.94]	0.101 (0.126) [0.80]				–0.0022
4	–0.065 (0.057) [–1.15]	–0.118 (0.126) [–0.94]	0.101 (0.127) [0.796]	0.0005 (0.063) [0.009]			–0.0098
5	–0.065 (0.057) [–1.155]	–0.119 (0.127) [–0.94]	0.102 (0.128) [0.798]	0.0009 (0.063) [0.015]	0.009 (0.066) [0.132]		–0.0173
6	–0.062 (0.057) [–1.092]	–0.099 (0.126) [–0.789]	0.071 (0.128) [0.554]	0.008 (0.063) [0.132]	0.025 (0.066) [0.382]	0.091* (0.055) [1.653]	–0.0041

Notes: *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively; numbers in parentheses stand for the OLS standard errors; numbers in square brackets are t-statistics.

The results of the first differences models show that all but one coefficient are insignificant. The only exception is f_6 in the last model, where the *t*-test (just) confirms the significance at the ten percent level. For the remaining cases, the largest value of the test statistic is observed for f_1 and f_2. The first differences models imply that changes in forward rates, neither one period before nor in the further lagged periods, have some predictive ability with respect to the changes in the spot rate. It is, thus, generally difficult to draw some meaningful conclusion from the first differences models. This result, combined with very low values of R^2, can be attributed to the specification of the model: estimating variables in first differences does not allow making such a meaningful inference as it would be possible in the case of a model expressed in levels.

3.2.2 The Error Correction Model

While table 15 represents the parameter estimates for the long-run relationship among the series, it is also interesting to investigate the short-run dynamics which is captured by the ECM. As cointegration was found, according to the Granger representation theorem,[74] an ECM is a valid representation of the data. An ECM is set up for each of the six models. The results are reflected in table 17, where both the estimates of CIVs and adjustment coefficients are presented. The latter are of great interest, as their significance indicates the validity of the error correction representation of the data. As table 17 suggests, with the exception of the first model where the significance is at the five percent level, the adjustment parameters are significant at the one percent level.

Thus, the conclusion that the respective series are cointegrated is reinforced by the significant adjustment parameters. This means that for these models the spot rate responds to the deviations from the long-run value, i.e. the ECM works. However, the adjustment coefficient δ_1 has a "wrong" sign in models 1 through 4. It is expected to be negative. In the case where the value of r_1 is above its long-run value, the change in r_1 should be negative to compensate for the disequilibrium in the previous period. However, out of six models, negative signs were observed only in case of models 5 and 6. Thus, only for these two models the ECM would make sense.

[74] See Engle/Granger (1987).

Table 17: The Error-Correction Model, Sample 1995 – 2006

$$\Delta r_t = \delta_0 + \delta_1 \left(r_{t-1} - \sum_{i=1}^{I} \hat{\beta}_i f_{i,t-i} - \hat{\beta}_0 \right) + \varepsilon_t$$

Model	1	2	3	4	5	6
f_1	1.2316* (0.2241) [5.49]	1.662* (0.261) [6.36]	1.645* (0.268) [6.12]	3.050* (0.459) [6.64]	–3.99* (0.558) [–7.16]	–1.441* (0.186) [–7.76]
f_2		–0.476*** (0.265) [–1.79]	–0.554 (1.004) [–0.53]	–3.353*** (1.829) [–1.83]	2.25 (1.978) [1.13]	–0.024 (0.633) [–0.04]
f_3			0.0134 (1.008) [-0.013]	2.95* (1.817) [1.62]	–2.20 (1.958) [–1.12]	–0.189 (0.618) [–0.30]
f_4				–1.338** (0.486) [2.76]	1.204*** (0.607) [1.98]	0.393** (0.194) [2.03]
f_5					2.099* (0.65) [3.22]	0.84* (0.217) [3.686]
f_6						0.579* (0.203) [2.85]
β_0	–0.015 (0.0098) [–1.55]	–0.01093 (0.01393) [0.78]	–0.007 (0.014) [–0.05]	–0.0009 (0.027) [–0.034]	–0.0173 (0.031) [–0.55]	–0.021 (0.0118) [–1.79]
δ_1	0.0291** (0.0117) [2.49]	0.034* (0.0089 [3.874]	0.026* (0.01) [2.62]	0.0219* (0.0049) [4.51]	–0.023* (0.004) [-5.85]	–0.079* (0.01) [–6.43]
R^2 (adj.)	0.111	0.098	0.118	0.129	0.20	0.23
F-statistic	9.39	14.89	4.43	20.08	33.98	41.11

Notes: *,** and *** denote the rejection of the null hypothesis at the 1%, 5% and 10% level, respectively; numbers in parentheses stand for the maximum likelihood standard errors; numbers in square brackets are *t*-statistics.

Finally, we consider the adjusted coefficients of determination in order to examine the goodness of fit of our models. The results are also not in favor of the forecast ability of forward rates. The adjusted R^2 of the first model only slightly exceeds 11 percent, for the second model this value is even lower, 9.8 percent. Then, starting from model 3 the adjusted R^2 increases gradually, achieving its highest level at 23 percent for model 6. The most significant increase by 7.1 percent in the coefficient of determination occurs when we move from model 4 to model 5, i.e. the inclusion of f_5 considerably improves the goodness of fit of the model.

3.2.3 Predictions

Although the results regarding the significance of coefficients and the goodness of fit are not very promising, the variables in all models are cointegrated and the error correction representation is valid for some of them. It is, therefore, of crucial interest, whether the fact that cointegration is present can help to improve forecasts. The estimated parameters for the sample 1995 to 2006 will be used to make forecasts for the period November 2006 to October 2007. To assess forecasting performance of our models, we compare the forecasts from the cointegration equations with the naive model which uses past period value of r_1 to make a forecast. The main purpose of this analysis is to measure the accuracy of the forecasts, which can be done with the help of various indicators. One of them is the mean absolute error (MAE), which can be computed in the following way:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|r_t - \hat{r}\right|, \tag{44}$$

where r_t and $\hat{r}_t$ denote the true and the forecasted values, respectively. Often a root mean squared error (RMSE) is employed:[75]

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(r_t - \hat{r}_t\right)^2}. \tag{45}$$

Another statistics which is commonly applied to test how closely the obtained forecasts match the data is the Theil's inequality coefficient, denoted as U:[76]

[75] See Pindyck/Rubinfeld (1998), p. 210.
[76] See ibid.

$$U = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(r_t - \hat{r}_t)^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} r_t^2} + \sqrt{\frac{1}{N}\sum_{i=1}^{N} \hat{r}_t^2}} \tag{46}$$

The numerator of equation (46) is represented by the RMSE; Theil's inequality coefficient lies between zero and one. A value of zero indicates a perfect match between the actual and forecasted values. For all other measures, a smaller value is desirable. The forecasting performance of the considered models is specified in table 18.

Table 18: Forecasting Performance

Model	MAE	RMSE	U
Naive	0,0104	0,0105	0,0555
r_1, f_1	0,0258	0,0263	0,1851
r_1, f_1, f_2	0,0280	0,0286	0,2114
r_1, f_1, f_2, f_3	0,0268	0,0275	0,1972
r_1, f_1, f_2, f_3, f_4	0,0343	0,0356	0,2959
$r_1, f_1, f_2, f_3, f_4, f_5$	0,0113	0,0475	0,0584
$r_1, f_1, f_2, f_3, f_4, f_5, f_6$	0,0061	0,0073	0,0368

Out of six estimated models, only model 6, which involves all six lagged forward rates, was able to beat the naive model according to all forecast accuracy measures. Model 5 is the second-best model according to the MAE and Theil's coefficient but not with respect to the RMSE which identifies model 1 as the second-best model. The model providing the worst forecast is model 4 according to the MAE and Theil's coefficient and model 5 according to RMSE. Thus, past forward rates exhibit rather poor predictive power with respect to the future one-year spot rate. Models including one to four lagged forward rates have no forecasting power at all, whereas for the model including six forward rates there seems to be some forecasting ability as it outperforms a naive model.

3.2.4 Summary

The pure expectations hypothesis asserts that long-term interest rates represent the average of the current and future expected short-term interest rates. Forward rates, which are derived from the term structure, are considered to be unbiased predictors of the future spot rates. According to this theory, bonds of different maturities are regarded as perfect substitutes. The pure expectations theory implies that bonds of different maturities, if held for an identical period of time should have the same holding period return. There exist, however, trading strategies aiming at gaining profits from the differences in holding period returns known as yield curve trading strategies. For example, the Riding the Yield Curve strategy involves buying a fixed income instrument with a maturity longer than the investor's holding period and selling it prior to maturity. The idea behind this fixed income strategy is to make use of the fact that the price of the bond increases, due to the fact that the interest rate on the remaining maturity is lower. However, if the pure EH holds, there should be no difference between the rate of return from riding and holding an instrument whose maturity matches the desired holding period. Another yield curve trading strategy is constructed through financing a long-term asset with a short-term liability. This strategy is widely applied by banks and is referred to as the Rolling Down the Yield Curve strategy.

A prerequisite for the success of both strategies described above is that the term structure of interest rates is upward-sloping and retains its shape in the future. A parallel shift or flattening out of the yield curve would result in a decline of the profits from the strategies. Because of a sharp shift of the term structure, the short-term interest rates could increase leading to losses on the strategies. According to the pure expectations theory, an upward-sloping yield curve indicates that future short-term interest rates are expected to rise. If normal term structure prevails, forward rates lie above the spot rates. Empirical results presented above, however, indicate that forward rates exhibit very low ability to explain the spot rates. Looking at the six models presented above, one can infer that, although some particular forward rates are significant, the general result does not confirm forward rates being unbiased predictors of the future spot rates. Regarding the empirical finding on their predictive ability, an upward-sloping yield curve does not indicate that spot rates will necessarily rise in the future. Consequently, strategies based on a stable, upward-sloping yield curve could be profitable. This aspect will be considered in details in the next chapter.

4. Yield Curve Trading Strategies

This chapter aims to present two commonly applied yield curve trading strategies and test their performance. In section 4.1 fixed income strategies are introduced and their classification is provided. In section 4.2 the main empirical findings of previous research are presented. The next sections 4.3 and 4.4 are devoted to the main characteristics of the RDYC strategy and the RYC strategy, respectively. Section 4.5 contains a detailed description of different performance indicators as well as the results of the performance evaluation analysis. The chapter finishes with section 4.6, which covers the regulatory framework and addresses the specific risks of the considered strategies.

4.1 Types of Strategies

Strategies based on the yield curve belong to the class of active portfolio strategies. In general, fixed income portfolio strategies can be roughly divided into active and passive strategies. The latter do not involve forecasting of factors that influence the performance of assets. An example of such a strategy would be investing in a portfolio that closely follows a broadly diversified bond index or a BH strategy involving buying an asset and holding it till maturity. Passive strategies are based on the postulate of the efficient market hypothesis[77] that markets are efficient and it is, therefore, not possible to outperform the market. In contrast, advocates of active portfolio strategies share the view that the markets are not completely efficient and, thus, seek to take advantage of a special forecasting ability of the portfolio manager. Passive strategies that do not require any analytical skills or forecasting abilities often serve as benchmarks for the performance of actively managed portfolios.[78]

Among the great variety of active fixed income strategies, one can distinguish between strategies based on market timing and strategies that attempt to identify mispriced securities. Yield curve trading strategies refer to strategies based on market timing, i.e. on the interest rate predictions. Two further subcategories of yield curve trading strategies can be distinguished:

[77] The weak form of the efficient market hypothesis in our framework states that it is not possible to predict future interest rates based on the publicly available information.
[78] See Martellini/Priaulet/Priaulet (2010), pp. 211-212.

strategies based on the specific changes of the yield curve and strategies which assume the yield curve shape will not change at all.[79]

4.1.1 Strategies Based on the Specific Changes of the Yield curve

Strategies that are based on anticipated changes in the yield curve shape are characterized by three main parameters: level, slope, and curvature.[80] Although each of these parameters or all of them together may change, the most common types of changes in the yield curve shape include upward and downward parallel shifts, twists and butterfly shifts.[81] Parallel shifts occur when interest rates change by the same amount across all maturities. Twists include changes in the spread between interest rates of different maturities and lead to either flattering out or steepening of the yield curve. Finally, a change in the curvature of the yield curve is referred to as a butterfly shift.[82]

Depending on the anticipated changes in the yield curve, bond portfolios are constructed in such a way, that the maturity of the included fixed income instruments perfectly matches these future changes. This class of strategies attempts to choose the duration of the bond portfolio that would match future interest rate changes. The main types of these strategies include *the barbell, the bullet and the ladder strategy*, although various combinations of them are also possible. The *bullet strategy* is constructed in such a way, that maturities of bonds are concentrated on one particular point of the yield curve. According to the *barbell strategy*, the maturities of the bonds are concentrated in two opposite ends of the yield curve; i.e. investing 50 percent of the portfolio value into a six-month maturity bond and another 50 percent into a 30-month maturity bond. Finally, the *ladder strategy* invests equal amounts into bonds of a range of maturities. Each of these strategies will perform differently, depending on the future changes in the slope, level and curvature of the yield curve.[83]

[79] See Martellini et al. (2003), p. 233.
[80] See Lettermann/Scheinkman (1991), pp. 54-61.
[81] See Fabozzi (2004), pp. 424-432.
[82] According to Jones (1991), these types of shifts do not occur independently. He analyzed the changes in the shape of the yield curve from 1979 to 1990 and came to a conclusion that downward shift with simultaneous steepening of the yield curve and an upward shift combined with a flattening of the yield curve were the most common to occur. The results obtained by Mann and Ramanlal (1997) support these findings.
[83] Although studies regarding the performance of these strategies are scarce, Man and Ramanlal (1997) test the performance of the barbell and the bullet portfolios. The also examined different maturities and reported that in the case of downward shifts of the yield curve short-maturity bullet portfolios performed better than short-term barbell portfolios. For upward shifts in the yield curve, the opposite was the case.

A very popular active fixed-income strategy is a combination of the barbell and the bullet strategy and is referred to as *butterfly strategy*. A standard butterfly is set up without having any initial costs of financing and zero duration, which makes the portfolio insensitive to small parallel shifts in the yield curve. A butterfly then involves a barbell consisting of a long term and a short-term bond and a bullet including a medium-term bond. Apart from the standard strategy, there exist several other butterfly trades that are not necessarily cash-neutral, namely fifty-fifty weighting, maturity weighting and duration weighting.[84]

Strategies based on the changes in the interest rate level assume that the only parameter of the yield curve that will change is the level of the curve, i.e. only parallel shifts may occur. If such changes are anticipated, the duration of the portfolio is adjusted to account for these changes. In case that interest rates are expected to increase, either the duration of the portfolio will be shortened or short-term bonds will be held till maturity and then "rolled over" at higher rates. The latter is referred to as *rollover strategy*. In contrast, if the portfolio manager anticipates a decline in interest rates, he or she will lengthen the duration of the portfolio.[85]

An additional yield curve trading strategy worth mentioning is a trading technique based on the *mean reversion* of the level, slope and curvature of the yield curve. If one of the three components lies above (below) its historical average level, the expectation is that in the future the level, slope or the curvature of the yield curve would fall (rise) in the direction of their historical average level.

4.1.2 Strategies Based on a Stable Yield Curve

Strategies based on the prediction that the yield curve remains stable are the *Riding the Yield Curve* strategy and the *Rolling Down the Yield Curve* strategy.[86]

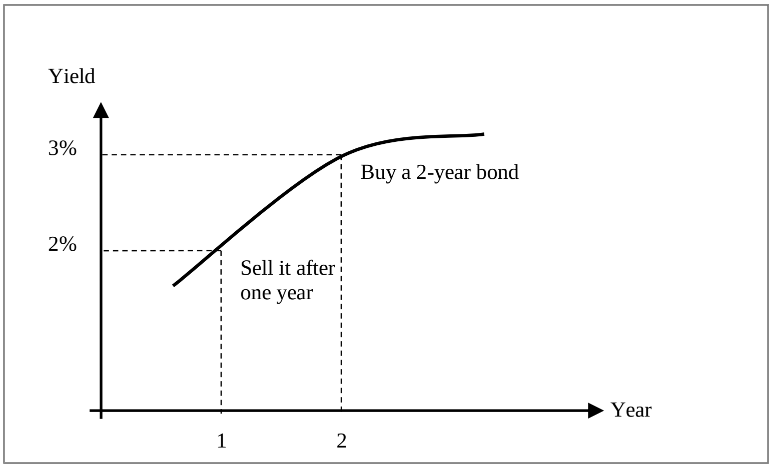
Riding the Yield Curve

Riding the Yield Curve is a popular way for a fixed income portfolio manager to enhance returns, which involves buying fixed-income instruments whose maturity is longer than the

[84] See Martinelli et al. (2002), pp. 9-13.
[85] See Martinelli et al. (2003), pp. 236-237.
[86] Although these two terms are sometimes used interchangeably, in this thesis it will be strictly distinguished between these two strategies.

investor's preferred holding period and selling them in the secondary market before maturity. An investor who would like to place his funds for some particular period may purchase a fixed-income security with maturity identical to this period; i.e. he or she can pursue a BH strategy. Alternatively, he or she could purchase a fixed-income security with maturity exceeding the desired holding period. After the intended investment horizon has been reached, the bond will be sold before it matures. If the yield curve remains unchanged, gains based on the falling bond yields as the bond's maturity decreases, can be made. For example, if the investor's holding period is one year, he or she could buy a two-year maturity bond and sell it after one year. An example of such a strategy, depicted in figure 8, shows that at the point of initiation of the strategy, a one-year and a two-year bond yield two percent and three percent, respectively. After one year, the maturity of the two-year bond purchased one year ago decreases and comprises only one year. If the term structure has not changed, the yield on a one-year bond would comprise only two percent now and, therefore, the bond can be sold at a higher price. In this situation, the rate of return on the RYC strategy is higher than the return the investor would have received from buying a security with maturity matching his holding period. The rate of return on the RYC strategy can be computed in several steps.[87] First of all, the price of a zero coupon bond with m years to maturity and a face value of 1,000 can be computed in the following way:[88]

Figure 8: RYC Strategy



$$P_t^m = \frac{1000}{\left(1+r_t^m\right)^m}. \tag{47}$$

[87] See Alexander/Sharpe/Bailey (2001), pp. 557-558.
[88] See Bieri/Chincarini (2005), pp. 7-11.

After *i* years, the bond is sold at the following price:

$$P_{t+i}^{m-i} = \frac{1000}{\left(1+r_{t+i}^{m-i}\right)^{m-i}}, \tag{48}$$

where r_{t+i}^{m-i} is the yield on the bond of maturity *m* – *i* available in period *t*+*i*. The rate of return on the strategy can be obtained by dividing equation (48) through equation (47):

$$r^{RYC} = \frac{\left(1+r_t^m\right)^m}{\left(1+r_t^{m-i}\right)^{m-i}} - 1. \tag{49}$$

Denoting r_t^i as the yield on a zero-coupon bond with *i* years to maturity, the excess return on the RYC strategy over a BH strategy can be represented in the following way:

$$exr^{RYC} = \left(\frac{\left(1+r_t^m\right)^m}{\left(1+r_{t+i}^{m-i}\right)^{m-i}} - 1\right) - \left(\left(1+r_t^i\right)^i - 1\right). \tag{50}$$

The forward rate for an investment starting in *t*+*i* and lasting *m* – *i* periods is given by:

$$f_t^{t+i,t+m} = \frac{\left(1+r_t^m\right)^m}{\left(1+r_t^i\right)^i} - 1. \tag{51}$$

The above expression for the forward rate can be inserted into equation (50) to replace the interest rate r_{t+i}^{m-i}:

$$exr^{RYC} = \frac{\left(1+r_t^m\right)^m}{\left(1+\left(\frac{\left(1+r_t^m\right)^m}{\left(1+r_t^i\right)^i} - 1\right)\right)} - 1 - \left(\left(1+r_t^i\right)^i - 1\right). \tag{52}$$

Rearranging the terms, the above equation can be simplified to:

$$exr^{RYC} = \left((1+r_t^i)^i - 1\right) - \left((1+r_t^i)^i - 1\right) \tag{53}$$

Thus, in case that the forward rate is used instead of the future spot interest rate, the excess return of the RYC strategy equals zero. For the above example of the RYC strategy using a two-year bond and holding it for two periods, the excess return would be computed in the following way:

$$exr^{RYC} = \left(\frac{\left(1+r_t^2\right)^2}{\left(1+r_{t+1}^1\right)^1} - 1\right) - \left(\left(1+r_t^1\right)^1 - 1\right) \tag{54}$$

Thus, the steeper the yield curve, the higher the gain from the strategy. In addition, the lower the short-term rate at $t+i$, the more profitable is the strategy. The capital gains from the sale of the long-term security are based on the lower remaining maturity of the instrument and the higher interest rate of the longer-maturity instrument.

A prerequisite for the success of the strategy is the normal term structure of interest rates. The main risk factor of the strategy is that the term structure will not remain stable and the short term rates could rise significantly. In this case, investors would have to sell bonds at a lower price and could possibly encounter losses on the strategy. Alternatively, the RYC strategy will yield an even higher rate of return, should the short-term interest rates fall by the time the long-term instrument is sold. The following table 19 provides an illustration of the returns on RYC strategy using different scenarios for the future short-term interest rate.

Table 19: Returns of the RYC Strategy under Different Scenarios

	Term structure at *t*		Short-term rate at *t*+1				
	r^2=4%	r^1=3%	2%	3%	4%	5%	6%
r_{RYC} BH Difference			6.04% 3% 3.04%	5.01% 3% 2.01%	4% 3% 1%	3.01% 3% 0%	2.04% 3% – 0.96%

Thus, if the interest rate on the short-term instrument stays the same or declines, the strategy yields five percent and six percent, respectively; this is more than the BH strategy returns. In the case that the interest rate rises and constitutes more than two percent, the investment starts to make losses and underperforms the BH strategy. In order to assess the riskiness of the strategy, the breakeven interest rate can be calculated. It shows, by which extent the interest rate can rise before the RYC strategy becomes just as profitable as the BH strategy. The breakeven rate $r_{BE,t}^{m-i}$ can be computed with the help of equation (50):

$$\left(\frac{\left(1+r_t^m\right)^m}{\left(1+r_{BE,t}^{m-i}\right)^{m-i}} - 1 \right) - \left(\left(1+r_t^i\right)^i - 1 \right) = 0. \tag{55}$$

From equation (55), the breakeven rate can be computed in the following way:

$$r_{BE,t}^{m-i} = \left(\frac{\left(1+r_t^m\right)^m}{\left(1+r_t^i\right)^i} \right)^{\frac{1}{m-i}} - 1. \tag{56}$$

In case of trading techniques with holding a two-year bond for one year, the breakeven rate will be represented by the following equation:

$$r_{BE,t}^{1} = \left(\frac{\left(1+r_t^2\right)^2}{\left(1+r_t^1\right)^1} \right) - 1. \tag{57}$$

Thus, the breakeven rate on a riding strategy using a m-year bond with i holding period is equivalent to the forward rate for an investment starting in t+1 and lasting $m - i$ periods. Investors may choose to implement the RYC strategy all the time, or they may apply different rules regarding when to pursue the strategy. Such rules are referred to as filter rules. The simplest filter rule is to ride when the yield curve is positively-sloped. The breakeven rate can also be applied in order to establish a filter rule. Such a filter rule could be, for instance, that the ride is executed only when there is a positive *cushion*, i.e. the difference between the breakeven rate and the short-term interest rate for $m - i$ periods at time t:

$$CU = r_{BE,t}^{m-i} - r_t^{m-i}. \tag{58}$$

In addition, the filter rule may be even more restrictive and specify that the cushion is not only positive, but comprises some particular percentage.

Rolling Down the Yield Curve

Similarly to the RYC strategy, the main idea of the *Rolling Down the Yield Curve* is also to exploit the upward-sloping yield curve, i.e. the fact that long-term interest rates usually lie above the short-term interest rates. The essence of this trading technique is to borrow funds at the short-term interest rate in order to invest them into a long-term asset. This strategy is widely applied in bond portfolio management and is a cash-neutral strategy, i.e. it does not require an initial investment. For example, an investor could purchase a two-year bond and finance his position through rolling over two one-year bonds. The rate of return on this strategy which invests in one *m*-years maturity instrument and finances it through rolling over *m* instruments with the maturity of one year can, thus, be computed in the following way:

$$r_{RDYC}^{m} = (1+r_t^m)^m - \left(\left(1+r_t^1\right)\cdot\left(1+r_{t+i}^1\right)\cdot\ldots\cdot\left(1+r_{t+m-1}^1\right)\right). \tag{59}$$

The RDYC strategy is equivalent to the RYC strategy, where the purchase of a long-term security is financed through the sale of a short-term instrument with the maturity $m-i$.

In the case that the strategy is implemented through purchasing a two-year instrument and financing it with two one-year instruments, the rate of return on such a strategy would look in the following way:

$$r_{RDYC}^{2} = (1+r_t^2)^2 - \left(1+r_t^1\right)\cdot\left(1+r_{t+1}^1\right). \tag{60}$$

If *m* is longer than two years, investors may also choose to use those instruments having a maturity longer than one year. For example, if the intended horizon constitutes three years, then the possibilities are:

1) purchase of a three-year instrument and finance it by rolling over three one-year instruments;

2) purchase of a three-year instrument and finance it through the sale of a two-year instrument in t and a one-year instrument in t+2;
3) purchase of a three-year instrument and finance it through the sale of a one-year instrument in t and a two-year instrument in t+1.

The excess return on the strategy using the first, the second and the third option is represented by equation (61), (62) and (63), respectively:

$$r_{RDYC}^{3} = \left(1 + r_t^3\right)^3 - \left(1 + r_t^1\right) \cdot \left(1 + r_{t+1}^1\right) \cdot \left(1 + r_{t+2}^1\right) \tag{61}$$

$$r_{RDYC}^{3} = \left(1 + r_t^3\right)^3 - \left(1 + r_t^2\right)^2 \cdot \left(1 + r_{t+2}^1\right) \tag{62}$$

$$r_{RDYC}^{3} = \left(1 + r_t^3\right)^3 - \left(1 + r_t^1\right)^1 \cdot \left(1 + r_{t+1}^2\right)^2. \tag{63}$$

As in the case of the RYC strategy, a prerequisite for the successful implementation of this strategy is that the upward-sloping yield curve does not change over time. In this case, the interest rate received from a long-term asset exceeds the interest rate paid on the short-term loan; i.e. a gain based on the long-short spread can be made. The risk of the strategy is that the shape of the yield curve will not remain constant and the short-term rates could rise. If the rise in the short-term rates is significant enough, the interest rate to pay on a short-term liability could exceed the interest rate received on a long-term asset and the strategy would suffer losses. After one year, the investor would face a higher one-year interest rate to finance his position. The breakeven rate for the strategy based on equation (60) may be computed as follows:

$$\begin{gathered} \left(1 + r_t^2\right)^2 - \left(1 + r_t^1\right) \cdot \left(1 + r_{BE,t+1}^1\right) = 0, \\ \Rightarrow \quad r_{BE,t+1}^1 = \frac{\left(1 + r_t^2\right)^2}{\left(1 + r_t^1\right)} - 1. \end{gathered} \tag{64}$$

As in case of the RYC strategy, different filter rules regarding when to pursue the RDYC strategy may be applied in attempt to enhance the returns from the strategy or to reduce its risk. The RYC strategy can be implemented in a variety of ways, using instruments with different maturities and holding them during different holding periods. Both of the above described techniques should not constantly persist. However, empirical evidence suggests some support for the effectiveness of this strategy. The existing empirical evidence regarding the strategy is given in section 4.2.

4.2 Previous Empirical Findings

Despite the variety of yield curve trading strategies, there exists only limited empirical evidence on their performance. Several studies which examined strategies based on a stable yield curve concentrated on the RYC and were predominantly based on US data. It was common for nearly all considered studies to compute the excess returns to the riding the yield curve strategy over the buy-and hold strategy.

Although the possibility to gain profits exploiting the RYC strategy has drawn the attention of both theorists and practitioners already in the 1970s,[89] the first attempt to empirically examine the riding the yield curve strategy was made by Osteryoung, Roberts and McCarty (1980). They investigated the US treasury bills market from January, 1976, to December, 1978, for maturities from seven to 182 days and a variety of holding periods. In this study the authors were the first to confirm that the RYC strategy provides higher rates of return than the simple buy-and-hold strategy. For almost all examined holding periods, the excess returns from the RYC strategy were statistically significant at the 95 percent level.

In the following study of Dyl and Joehnk (1981) over a longer horizon covering the period 1970 – 1975, the authors examined the returns from the RYC strategy in the US market for maturities up to 20 weeks and holding periods of four, eight, 12 and 16 weeks. Their overall result is that the returns from the RYC strategy are higher than those obtained from the BH strategy with the returns increasing with the length of the holding period. The strategy was pursued not all the time, but conditioned on different filter rules. The authors applied filter rules based on the breakeven rate, i.e. the rate at which the riding strategy just becomes unprofitable. They created a so called "margin of safety", which represents the difference between the break-even rate and the interest rate on a bond of the desired holding period. In general, the excess returns tend to increase with the longer holding period and higher margin

[89] Some early textbooks describing riding the yield curve strategy include Freund (1970), pp. 66-67 and Darst (1975), pp. 290-295. In the practitioner's literature, riding the yield curve was mentioned by De Leonardis (1966), pp.48-53 and later by Van Horne (1974), p. 361.

of safety. However, only in a few cases, namely for the highest margin of safety and longest holding period, statistically significant returns could be achieved.

Grieves and Marcus (1992) proved that previously obtained results on riding the yield curve also hold true during a significantly longer period ranging from 1949 to 1988. They performed three-month rides with three-, six- and 12-month treasury bonds. Compared to the BH strategy, RYC, conditioned on the same filter rules as those used in Dyl and Joehnk (1981), produced higher returns for almost all considered maturities and margins of safety. According to the authors, a three-month ride with six-month T-bills proved to be especially successful across the riding strategies. For this type of ride and over several subperiods and filter rules, they found that the RYC strategy dominates the BH strategy by the first degree stochastic dominance criterion. However, this result did not hold for the whole sample and for maturities exceeding six months.

The following study of Chandy and Hsueh (1995) uses bonds with up to 13 weeks maturity for the period of 1981 – 1985 to test unconditional riding as well as riding based on a positive slope of the term structure. They report that all tested riding strategies, conditional as well as unconditional, provide positive excess returns compared to the BH strategy. The returns tend to rise with the maturity of the selected instrument, but not with the holding period. However, significance tests revealed that none of these excess returns is statistically significant. Therefore, the strategy should not be pursued by the investors, who are better off just buying an instrument of the desired holding horizon and holding it till maturity. Although riding instruments with longer maturities yield higher excess returns, these returns were accompanied by an increase in the standard deviation of the excess returns.

Whereas previous research has concentrated on very short holding periods and maturities only up to 12 months, Pelaez (1997) considers riding the two-year US treasury securities for one year. He performs rides whenever the term structure was upward-sloping over the period 1959 – 1993. He demonstrates that the strategy also works with longer maturity instruments, as the mean return from the riding strategy turns out to be 75 basis points higher as opposed to the BH strategy. However, he notes that the risk associated with the RYC strategy, measured by the variance of the returns, is twice as high as that of the BH strategy. Pelaez (1997) interprets his findings as evidence in favor of the liquidity premium theory of the term structure.

A more extensive study considering several countries and maturities was provided by Ang, Alles and Allen (1998). Their data set covers the period from 1985 to 1996 and includes instruments of maturities ranging from several weeks to two years for Australia, the US, Canada and the UK. Similarly to the findings of Dyl and Joehnk (1981), their results are maturity-dependent, i.e. higher rates of return to the RYC strategy were documented for instruments of longer maturities. In particular, in case of one-year holding period, the RYC strategy provides significant positive excess returns compared to the BH strategy for all considered countries. In contrast, for holding periods not exceeding 13 weeks, the RYC strategy proves to be inferior to the BH strategy in most cases.

However, the results of the previous research concerning the ability of the filter rule to enhance returns could not be confirmed. Filters based on the positively sloped yield curve as well as filters based on different margins of safety proved to be useless for all considered strategies. The authors also studied the influence of the transaction costs on the strategy's return and found out that negative excess returns have resulted when the transaction costs were set out to be more than 0.125 percent. However, the transaction costs did not have such a strong influence on returns in case of longer holding periods. In addition, the authors checked the riding strategies for the presence of the stochastic dominance over the corresponding BH strategy using the first degree stochastic dominance criterion. As opposed to the finding of Grieves and Marcus (1992), who found the RYC strategy to dominate the BH strategy in terms of the first degree stochastic dominance over several subperiods, they could not confirm stochastic dominance for their data set. The only exception was riding the US six-month bills, where the first degree stochastic dominance could be confirmed.

Grieves et al. (1999) consider the US market over the period 1987 – 1997 for maturities up to 12 months and try several riding strategies, for holding periods of three and six months, conditional as well as unconditional on a filter rule. They report positive excess returns over the BH strategy for all considered rides. The authors were able to confirm some previous findings regarding higher returns for longer maturities. This, however, comes at the cost of additional risk, which is measured by the standard deviation. For conditional rides, three filter rules are employed: to ride when the slope is positive; to ride when the yield spread exceeds 0.15 percent and, finally, to ride when the yield spread is larger than 0.30 percent. Whereas the first filter rule did not yield higher excess returns compared to unconditional rides, the latter two restrictions resulted in a substantial increase in the strategy's returns. This, however, has dramatically reduced the number of rides, which constituted only four percent of all rides for some maturities.

Bieri and Chincarini (2005) checked the effectiveness of the riding strategy for both bond and swap markets in the US, the UK and Germany. Employing a wide range of maturities and holding periods, their study for the period 1973 – 2003 suggests that the RYC strategy is superior to the BH strategy in most cases. In addition, using longer maturity riding instruments resulted in greater returns for all countries. However, only for the UK market it was possible to enhance returns when riding with both longer maturities and holding periods. In the case of Germany and the US, the returns to the RYC strategy increased with longer maturity of the riding instrument, but not the holding period. In fact, the highest returns for these countries were obtained when combining long-maturity instruments with short holding periods. The performance of different riding strategies was evaluated using Sharpe ratios. On the risk-adjusted basis, the most beneficial strategy for Germany was to ride two-year bonds, the longest considered maturity.

In addition, Bieri and Chincarini (2005) tested, whether returns from the RYC strategy can be further increased when the riding is conditional on a filter rule. Among such filters the authors used a positive slope filter and a positive cushion filter. The latter indicates that the interest rates still can increase without that the riding strategy becomes unprofitable. For the US and Germany, the rides conditioned on positive slope of the yield curve did not perform better than when the strategy was pursued in any time. For the UK, a positive slope filter brought the improvement of the returns as opposed to the unconditioned ride, but this only occurred for holding periods below 12 months. In contrast, the RYC strategy based on a positive cushion filter was more successful and yielded returns significantly higher than the unconditioned riding strategy for the majority of countries and maturities. It is worth noting that the positive cushion filter rule was the most successful over the short holding horizons.

Finally, Mercer, Moore and Winters (2009), pointed out that, despite the obvious profitability of the riding strategy documented by the previous research, the returns from the strategy continue to persist. Based on the US T-bills market during the period 2001 – 2007, they have investigated whether these returns can be explained by the fact that the transactions necessary to implement the RYC strategy were not available. In this period, they found supporting evidence for the profitability of the riding strategy involving selling 182-days T-bills after 91 days. However, pursuing the RYC strategy would require selling off-the-run T-bills[90] after a 91-days holding period was reached. As off-the run T-bills are by far not as frequently traded as on-the-run T-bills, selling them to execute a riding transaction could be problematic. The authors analyzed the trading volume on a secondary market and came to a conclusion that although the RYC strategy would generate significant profits, it would be necessary to capture almost all existing trading volume in order to generate a gain equal to one million dollar. According to the authors, market limitations prevent market participants from exploiting the RYC strategy.

Thus, there exist only a limited number of studies concerning the riding strategies. Most of them concentrate on the US market and compare the return from the riding strategy with a benchmark, which is in all cases the BH strategy. These findings show that the RYC strategy provides positive excess returns over the BH strategy. Most of the studies document that excess returns tend to be higher when the strategy is implemented with longer maturities instruments. In several cases, excess returns from riding increased with the length of the holding period. The returns could be sometimes enhanced when some further restrictions in form of a filter rule were imposed on the strategy. However, such an increase in excess returns only occurred if some sophisticated filter rule was applied; a simple filter rule based on a positively-sloped yield curve did not prove to be effective.

Although these studies provide evidence that the RYC strategy yields excess returns compared to the BH strategy, the emphasis was made on the excess returns and not on the riskiness of the strategy. Performance on a risk-adjusted basis was not considered in most of the studies. In the following sections we consider the RDYC strategy and the RYC strategy using data for Germany. The success of the strategy is evaluated with the help of several performance measures. Then the risk-adjusted return of the strategy is compared with different benchmarks, including the stock market.

[90] A T-bill becomes off-the-run when a new T-bill of any maturity is issued.

4.3 Rolling Down the Yield Curve

Rolling down the yield curve involves borrowing funds at the short-term interest rate and investing them in long-term assets. From a bank's position, such a strategy belongs to the traditional loan-granting and deposit-taking activities. Generally, one can distinguish between the lending business and fee-based, market-side activities.

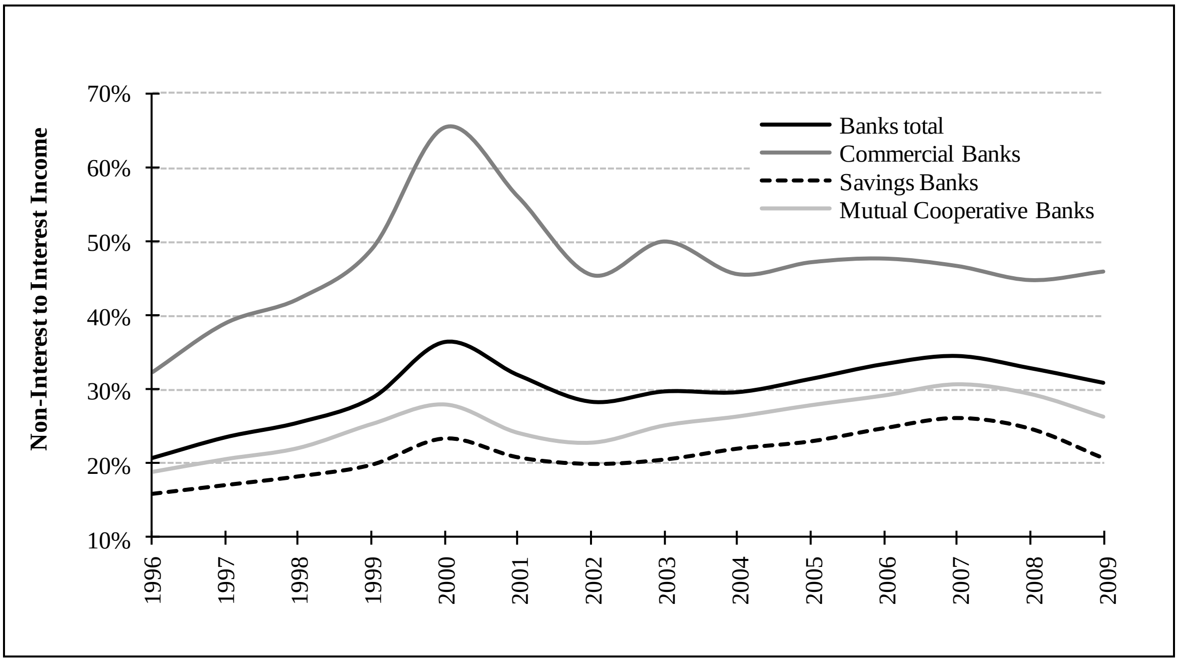
In Germany, most banks can be classified as universal banks, which are permitted to engage into a wide range of banking activities. However, there are some certain groups of banks where some specific source of income plays an important role. Depending on their legal form, universal banks can be further divided into three main categories: commercial, savings and mutual cooperative banks. This three-pillar structure is a characteristic feature of the German banking system. Savings banks are banks which are owned by the federal, state or local municipalities mainly concentrate on lending to individuals as well as small- and middle-sized enterprises whereas mutual cooperative and commercial banks are privately owned. Mutual cooperative banks comprise 60 percent of all banks in Germany at the end of 2008 and represent the largest banking group.[91] However, considering the amount of assets per banking sector, the savings banks sector shows the largest share of total assets, accounting for 35 percent of total assets in the German banking system. This significant amount is a remarkable feature of the German banking structure.[92] Figure 9 demonstrates the share of non-interest income to interest income for all banking sectors in the past decade.[93] This ratio reflects the relative importance of alternative income sources.

The savings banks sector exhibits the lowest ratio of non-interest income to interest income compared to the commercial and especially to the cooperative banks sector. In the past decade, the average share of non-interest income to interest income for savings banks was only 22.6 percent. Commercial banks, whose ratio on average comprised 50 percent in the past decade, experienced a shift towards fee-based activities, such as asset management and investment banking. In contrast, for savings banks traditional loan-granting and deposit-taking activities remain the primary source of income, although the importance of the fee-based activities slightly increased in the past decade.

---

[91] See Reichling/Afanasenko (2010), pp. 2-3.
[92] See Reichling/Afanasenko (2010), pp. 5-6.
[93] The calculations for his figure are performed on the basis of the various monthly reports of the German Central Bank; see German Central Bank (2000, 2004 and 2010).

Figure 9: Non-Interest Income to Interest Income for Different Banking Groups



Thus, the RDYC strategy reflects the lending business activities of banks and will be considered from the point of view of primary bank interest business. As opposed to it, the second main bank business, the banks' activities on the capital market, will be represented by the investment into the German Stock Index (DAX).[94] The main goal is then to compare these two types of bank activities based on the performance of the corresponding strategies. The following section provides the general description of the RDYC strategy and examines its excess returns in comparison to the DAX.

4.3.1 The Main Characteristics

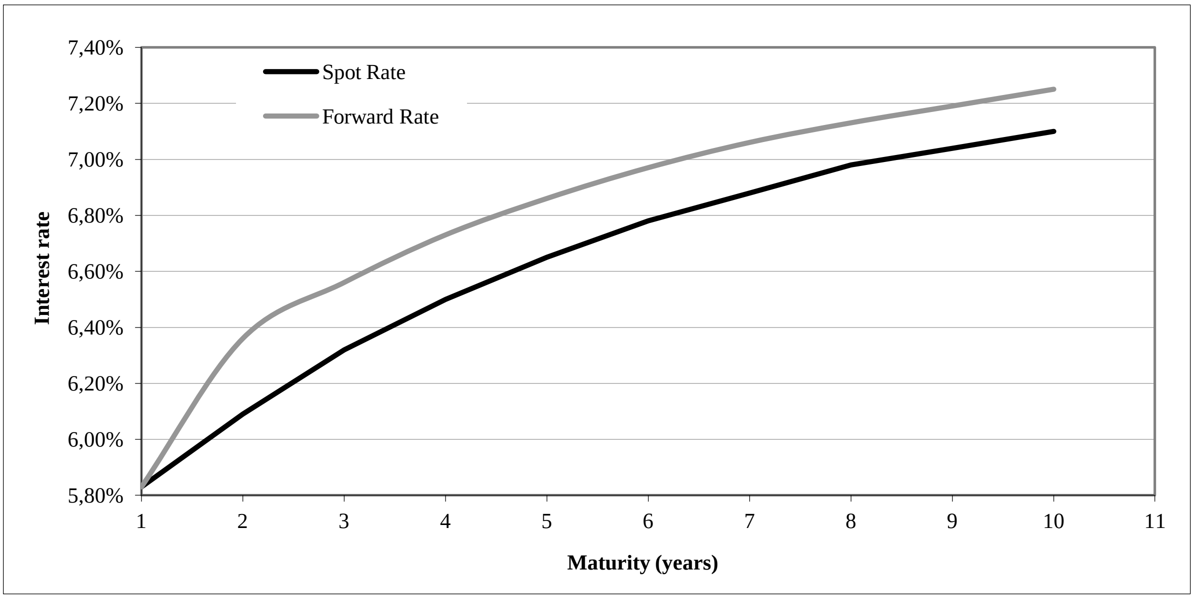
In the following we pursue the RDYC strategy and the RYC strategy. The analysis is based on the data set consisting of the German swap rates for the period ranging from 1972 to 2007. From the swap curve, the interest rates for a wide spectrum of maturities were calculated according to the equation (5). Figure 10 shows the average spot interest rates along with the corresponding average forward rates. On average, the term structure of interest rates in Germany for the period 1972 – 2007 was upward-sloping. Thus, necessary prerequisites for implementing yield curve trading strategies are met on the German market. One can observe that forward rates lie above the corresponding spot rates.

[94] In German: Deutscher Aktienindex (DAX).

Several types of the RDYC strategy are considered including the following riding periods: two-year ride, three-year ride, four-year ride and five-year ride. The strategies are denoted as $RDYC_m$, where m stands for the maturity of the long-term instrument being purchased. All long-term maturity instruments are financed through rolling over an appropriate number of one-year rates. For example, the two-year strategy involves purchase of a two-year instrument which is financed through the sale of a one-year instrument and, in one year, another one-year instrument. It would be, of course, possible to fund the purchase of a long-term asset through rolling over instruments with even lower maturity, for example, less than one year. However, we do not possess data for such short maturities and, therefore, do not consider such kind of strategies in this study.

In addition, every type of strategy is pursued in two different ways: riding all the time, independently of the prevailing term structure at the moment of initiation and riding using a filter rule. The latter is based on a simple principle: the strategy is implemented if the term structure is upward sloping, i.e. $r_t^m > r_t^{m-1}$; otherwise no investment strategy is implemented in the corresponding period. Using such a filter rule is motivated by the results of previous research, which could in some cases document higher returns using a filter rule.

Figure 10: Average Spot and Forward Rates from 1972 to 2007



An important difference between the two strategies is that the RDYC is a cash-neutral strategy whereas the RYC requires an initial investment to buy a respective fixed income instrument. Consequently, in case of the RDYC strategy the return is an excess return rather than rate of return, as it is a cash-neutral strategy where the invested amount equals zero. It is, therefore, more appropriate to speak about excess return, which is for the RDYC strategy measured as the difference between the interest rate on the instrument of maturity *m* and the interest rate on *m* one-year instruments, defined by equation (59). In order to compare the performance, the DAX is chosen as an equity investment benchmark. To insure the comparability of the cash-neutral RDYC investment strategy with the benchmark, it is assumed that the investment in DAX was financed by borrowing funds for one month.

4.3.2 Excess Returns

The excess returns of the strategy where the two-year investment is financed through two revolving one-year liabilities (RDYC 2) and the strategy involving a three-year investment (RDYC 3) are reflected in Figure 11. The excess returns of the four-year (RDYC 4) and five-year (RDYC 5) strategy as well as the DAX are plotted in Figure 12.

Figure 11: Excess Returns of the two- and thee-year strategies and the DAX

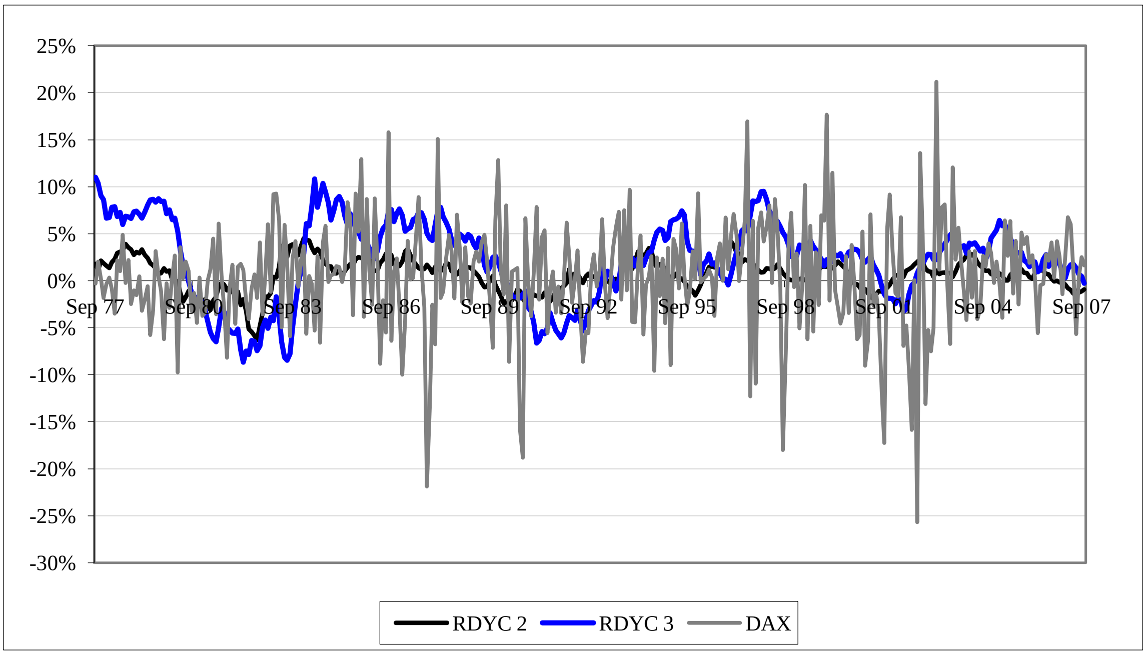
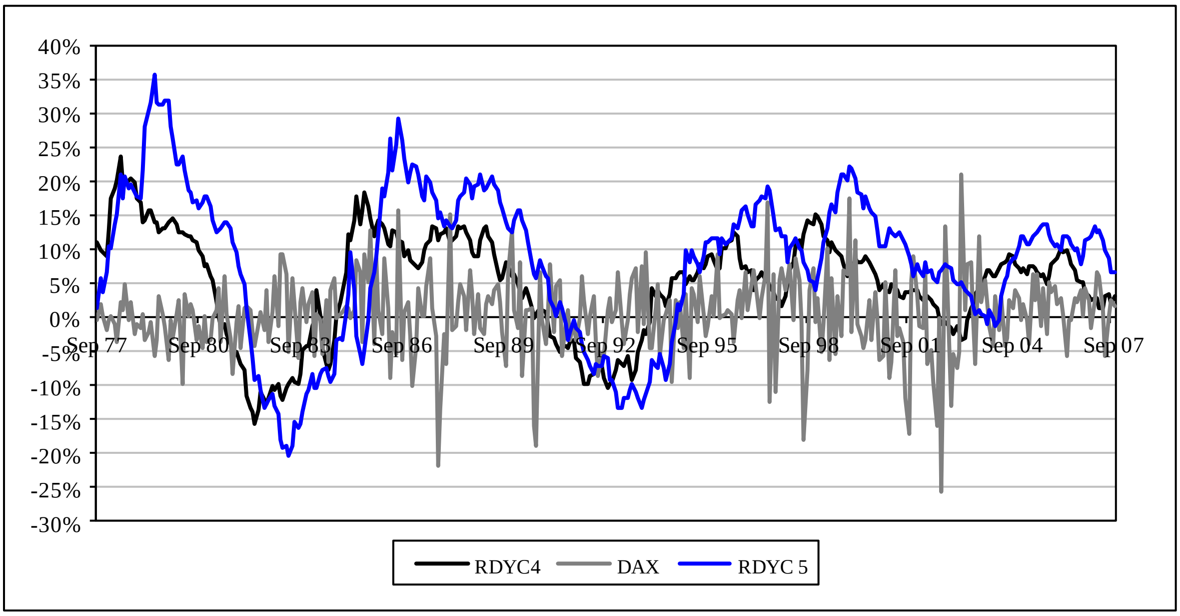


Figure 12: Excess Returns of the four- and five-year Strategies and the DAX



All returns were calculated without considering the transaction costs. In the case that the strategy is pursued by institutional investors through buying and selling corresponding fixed income instruments, the transaction costs are negligibly low. For example, prices for a bond transaction quoted by the German Stock Exchange (XETRA Group) constitute from 6.5 to 7.48 bps, depending of the volume of the order.[95] From the other side, if the RDYC strategy is implemented by means of loan-granting and deposit-taking, the costs are counted as an extra charge. If, finally, the strategy is implemented by means of swap transactions, there are no transaction costs. Thus, transaction costs are either non-existent or negligible and, therefore, are not considered in the following analysis. At the first glance, it is evident that the excess returns of the DAX are more volatile than the excess returns of the RDYC 2 and RDYC 3. However, the excess returns of both strategies with the longest maturities exhibit higher volatility than the DAX.

Table 20 reports some summary statistics of the RDYC strategy, both unconditional and using a filter rule. The latter strategies are marked with “F”. The success rate refers to the percentage of cases in which the corresponding strategy yielded positive excess returns. In addition, riding frequency was calculated for the RDYC conditional strategies. As table 20 shows, the mean excess return for all strategies is positive. For unconditional strategies, the lowest mean excess return of 0.79 percent per month was achieved by investing in a two-year bond and financing it through rolling over two one-year bonds. The mean excess return increases sharp-

[95] See German Stock Exchange (2011), p. 10.

ly with the maturity of the applied instrument and constitutes more than eight percent for the five-year instrument. All strategies have higher mean excess return than the DAX, except of the two-year strategy, whose return is slightly lower. The difference in the return obtained from the two-year and the five-year instrument constitutes more than seven percent. However, not only the excess return, but also the volatility is higher for longer maturity instruments.

Table 20: Summary Statistics of the RDYC Strategy

Strategy	Mean excess return	Std. Dev.	Minimum	Maximum	Success Rate	Frequency
			Unconditional			
RDYC 2	0.76%	1.85%	-6.30%	6.05%	71.36%	100%
RDYC 3	2.45%	4.42%	-8.66%	14.02%	74.35%	100%
RDYC 4	4.93%	7.63%	-15.56%	23.81%	76.47%	100%
RDYC 5	8.18%	11.26%	-20.46%	35.88%	77.35%	100%
DAX	0.42%	5.86%	-25.42%	21.38%	62.00%	100%
			Conditional			
RDYC 2F	0.69%	1.47%	-3.25%	6.06%	60.55%	81.5%
RDYC 3F	1.91%	3.99%	-9.68%	13.12%	55.18%	71.0%
RDYC 4F	3.37%	6.83%	-15.56%	23.81%	53.48%	70.6%
RDYC 5F	4.93%	10.04%	-20.46%	35.87%	51.10%	70.1%
DAX	0.42%	5.86%	-25.42%	21.38%	62.00%	-

Note: The number next to the name of the strategy denotes the number of years of the long-term investment; an "F" means that the strategy was based on a filter rule.

Thus, the five-year strategy also exhibits the highest volatility, which constitutes 11.26 percent. In contrast, the volatility of the shortest maturity instrument is only 1.85 percent. In addition, the success rate increases along with the instrument's maturity. The five-year strategy yielded positive excess returns in 77 percent of all cases whereas the two-year strategy yielded positive excess returns in 71 percent of all cases. Compared to unconditional strategies, investing in the DAX provides the lowest success rate of only 62 percent.

The applied filter rule led to the different frequency of riding for the strategies, as the filter rule became more restrictive with longer maturity. For the two-year strategy, the rule is $r_t^2 > r_t^1$; for the three-year strategy, it is $r_t^3 > r_t^2 > r_t^1$. Whereas for the two-year rolling strategy the filter gave a riding signal in more than 80 percent of all cases, the five-year strategy was implemented only in 70 percent of all cases. The filter rule reduces the frequency of riding by ten percent when we move from the two-year to the thee-year strategy, but makes almost no difference for the rest of the strategies.

Conditional strategies show lower returns and standard deviations than the strategies where no filter was applied. The mean excess returns lie in the range of 0.69 to 4.93 percent. Similarly to the unconditional strategies, the returns and standard deviations tend to increase with the maturity of the instrument. The five-year conditional strategy is characterized by an excess return of 4.93 percent and a volatility of ten percent. As in case of the unconditional strategies, the only strategy providing lower mean excess return than the DAX is the two-year strategy, whose return only comprises 0.69 percent. Also the success rates are lower for the conditional strategies and tend to decrease with the longer maturity. The two-year conditional strategy ended in positive excess returns in 60 percent of all cases, whereas for the five-year strategy it happened in only slightly more than the half of all cases. All considered strategies, both conditional and unconditional, exhibit higher excess returns than the DAX.

In addition, it is tested whether the excess returns of various strategies are significantly different from zero. For this purpose, two types of significance tests are performed. The first is a standard *t*-test; the second test was developed by Newey and West (1987) and is known as the Newey-West (NW) estimator. The NW test, although not as powerful as the *t*-test, stays consistent also in the presence of autocorrelation and heteroscedasticity.[96] The results of both tests that are reported in table 21.

Table 21: Significance Tests of the Excess Returns on the RDYC Strategy

Strategy	RDYC 2	RDYC 3	RDYC 4	RDYC 5
Unconditional				
t-statistic	8.208	10.885	12.489	13.988
p-value	0.000	0.000	0.000	0.000
NW test	3.594	4.623	5.235	5.837
p-value	0.0004	0.000	0.000	0.000
Conditional				
t-statistic	9.388	9.379	9.541	9.346
p-value	0.000	0.000	0.000	0.000
NW test	4.244	4.084	4.129	4.055
p-value	0.000	0.0001	0.000	0.0001

Note: The number next to the name of the strategy denotes the number of years of the long-term investment.

The *t*-test indicates that the excess returns of both conditional and unconditional RDYC strategies are significantly different from zero, as the *p*-values are extremely low. The NW test,

[96] See Brooks (2008), p. 152.

although yielding lower values of the statistic, confirms the findings of the *t*-test. Thus, the excess returns of both conditional and unconditional rolling strategies are significantly different from zero.

4.3.3 Excess Return - Volatility Profile

Figure 13 visualizes the excess return - volatility profiles of all considered strategies including the DAX. Already at the first glance, it is evident that the two-year and three-year strategies, RDYC 2 and RDYC 3, clearly dominate the DAX, as they simultaneously provide higher return and lower volatility. RDYC 4, although exhibits higher volatility than the DAX, is also characterized by a substantially higher excess return. Comparing conditional strategies with unconditional ones, the latter lie higher and to the right of the former ones, as they provide higher returns, but also exhibit higher standard deviations. However, the increase in return is generally higher than the increase in volatility. An additional insight is provided by figure 14, which compares the histograms of four riding strategies with that of the DAX.

Figure 13: Excess Return-Volatility Profile of the RDYC Strategies

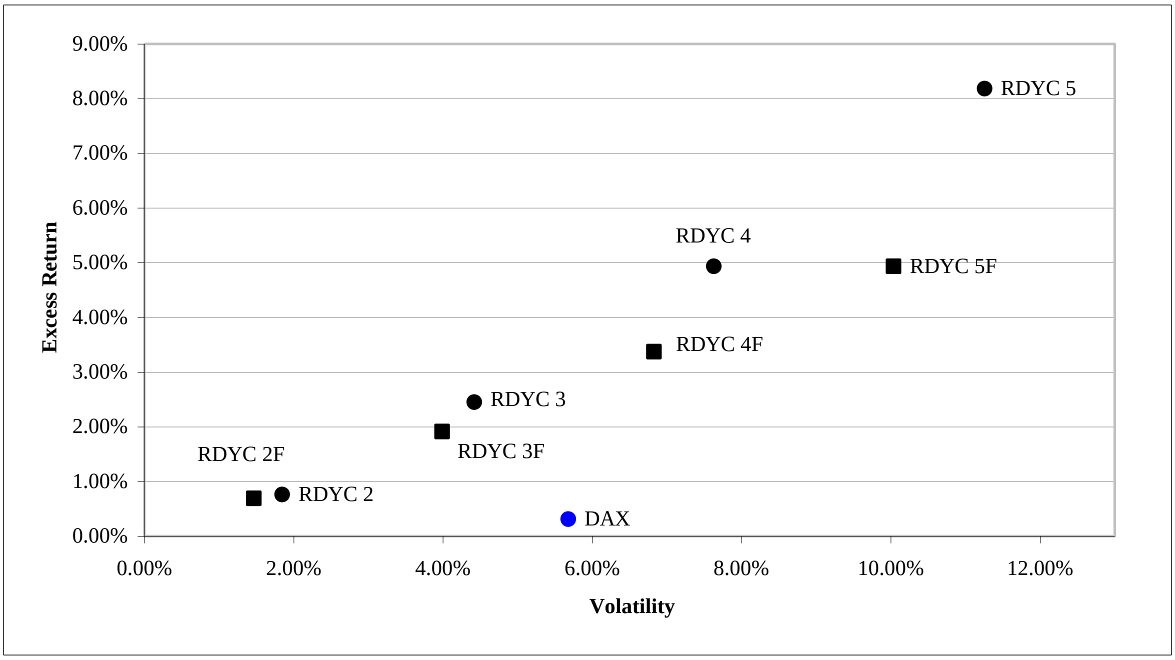
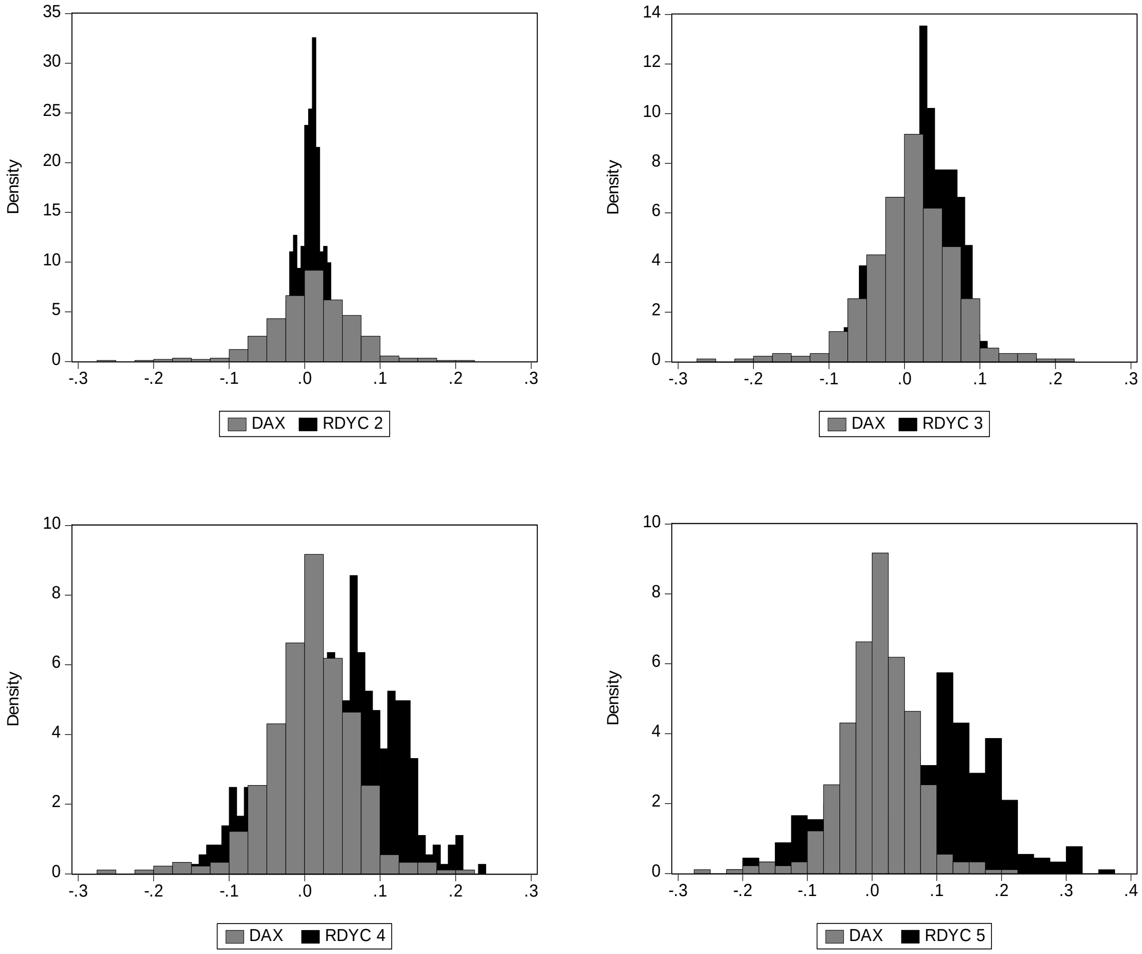


Figure 14: Histogram of the RDYC strategies and the DAX



4.4 Riding the Yield Curve

4.4.1 The Main Characteristics

The RYC strategy is constructed as a purchase of a two-year, three-year, four-year or five-year instrument and a sale of this instrument before maturity. If an interest rate is for *m* years, different rides are applied, varying the length of the holding period up to $m - 1$. Different strategies are denoted as $RYC_{m/i}$, denoting the RYC strategy with purchase of an *m* years maturity instrument and selling it after holding it for *i* years. For example, we purchase a five-year instrument and try four different strategies: selling after one, two, three and four years. In addition, it will be checked if a positive slope filter rule is able to enhance the returns from the RYC strategies. Similarly to the RDYC strategies, the strategy is implemented conditioned on the positive slope of the term structure in the considered period. If this condition is not fulfilled, the funds are just invested in the instrument of the desired holding period. Altogether, 28 strategies are considered: eight RYC strategies and 20 RDYC strategies. For the RYC strategy, the rate of return is defined by equation (49) and the excess return is normally computed as the difference between the rate of return and the interest rate that would prevail if an investor invested in accordance with his desired holding period.

4.4.2 Excess Returns

The excess returns of the RYC strategy are computed using a BH strategy as a benchmark, i.e. investing in a bond of the respective maturity. For example, if we ride a two-year bond for one year, the benchmark is an investment in a one-year bond. For a two-year ride of a three-year bond, the benchmark is an investment into a two-year bond at the interest rate which was prevailing at the time where the strategy was initiated. Figure 15 depicts the rates of return on one-year riding strategies with two- (RYC 2/1) and three-year (RYC 3/1) maturity instruments together with the BH strategy.

The statistics regarding the RYC strategies, which are performed through buying an *m* years maturity instrument and selling it after *i* years are summarized in table 22. The excess returns of the RYC strategies are all positive, i.e. the investors are better off performing a riding strategy rather than just buying an instrument and holding it till maturity. The returns lie in the range 0.70 percent for riding a two-year bond for one year to 4.07 percent in case of the riding strategy involving a five-year bond and holding it for three years. Standard deviations range from 1.72 percent to 6.54 percent for RYC 2/1 and RYC 5/2, respectively. Performing riding strategies with instruments of longer maturity generally yields higher excess returns. However, it also results in higher standard deviations.

One more observation concerns different holding periods within a particular maturity. Moving from a shorter holding horizon to a longer one leads to an increase in the excess return. This, though, does not happen in all cases. For example, riding a four-year bond for two years instead of one year yields an increase in the excess return from 1.79 percent to 2.73 percent. However, if the holding horizon constitutes three years, the excess return drops to 2.21 percent. Similarly, increasing the holding horizon of a five-year bond from one to first two, and then to three and four years leads to an increase in the excess return from 2.20 percent to 3.77 percent and then to 4.07 percent.

Figure 15: Rate of Return-Volatility Profile of the RYC Strategies

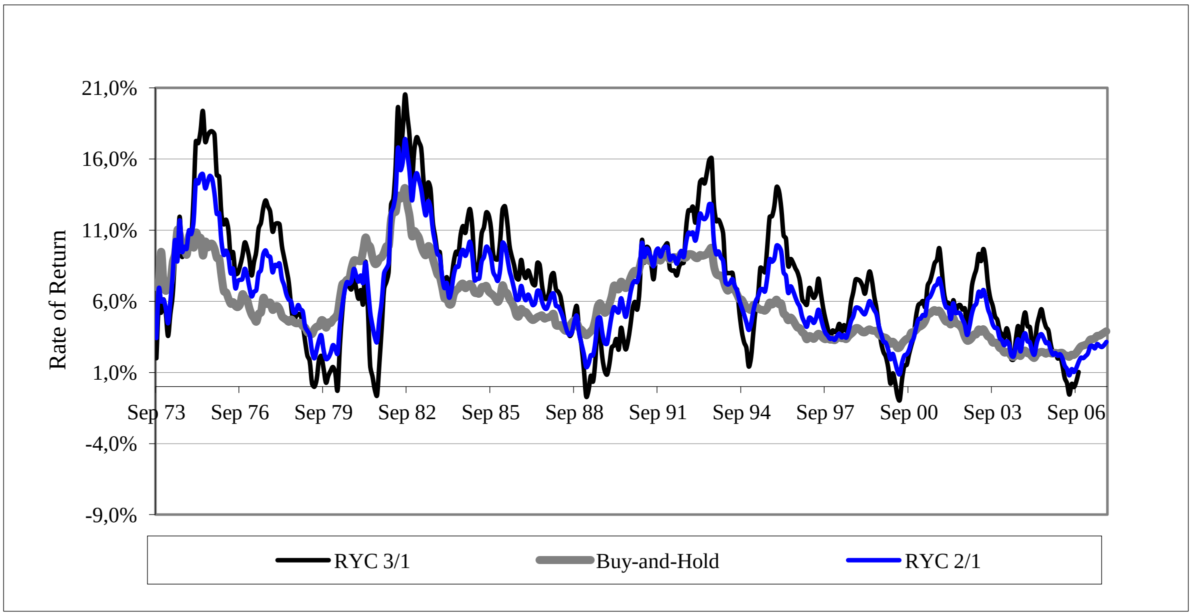


Table 22: Summary Statistics of the RYC Strategy

Strategy	Mean excess return	Std. Dev.	Minimum	Maximum	Success Rate
RYC 2/1	0.695%	1.72%	-5.53%	5.72%	69.27%
RYC 3/1	1.37%	3.24%	-9.26%	10.78%	68.34%
RYC 3/2	1.49%	2.63%	-6.11%	7.28%	71.86%
RYC 4/1	1.79%	4.45%	-12.54%	13.26%	66.34%
RYC 4/2	2.73%	4.79%	-10.55%	11.98%	71.11%
RYC 4/3	2.21%	3.15%	-7.44%	8.82%	76.68%
RYC 5/1	2.20%	5.50%	-15.37%	15.31%	67.32%
RYC 5/2	3.77%	6.54%	-14.73%	15.75%	71.61%
RYC 5/3	4.07%	5.79%	-13.10%	17.19%	76.94%
RYC 5/4	2.81%	3.46%	-7.63%	10.68%	80.21%

Note: The first number next to the name of the strategy denotes the maturity of a fixed income investment whereas the second number stands for the holding period.

However, for the last strategy it drops again and constitutes only 2.81 percent. As for the standard deviation within particular maturities, it is difficult to observe a clear pattern, although it tends to at first rise and then drop when switching to the longest possible holding period. Whereas the standard deviation constitutes 4.45 percent in case of riding a four-year bond for one year, holding it for two years yields a standard deviation of 4.79 percent; for three years – 3.15 percent. The success rate tends to increase with longer maturities and holding periods.

It is evident from table 23, which contains the results of the *t*-test as well as the NW test, that the excess returns of all riding strategies are significantly different from zero at a very high significance level. This is implied by the *p*-values, which are very low for all strategies.

Table 23: Significance Tests of the Excess Returns on the RDYC Strategy

Strategy	RYC 2/1	RYC 3/1	RYC 4/1	RYC 5/1	RYC 3/2	RYC 4/2	RYC 5/2	RYC 4/3	RYC 5/3
t-stat.	8.201	8.441	8.131	8.087	11.274	11.378	11.504	13.778	13.828
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NW test	3.582	3.678	3.540	3.525	4.573	4.802	4.865	5.752	5.784
p-value	0.0004	0.0003	0.0004	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000
t-stat.	9.388	9.549	9.603	9.582	8.709	9.608	9.813	14.046	8.865
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NW test	4.087	4.237	4.251	4.242	3.788	4.174	4.248	5.992	3.818
p-value	0.0001	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0002

Note: The first number next to the name of the strategy denotes the maturity of a fixed income investment whereas the second number stands for the holding period.

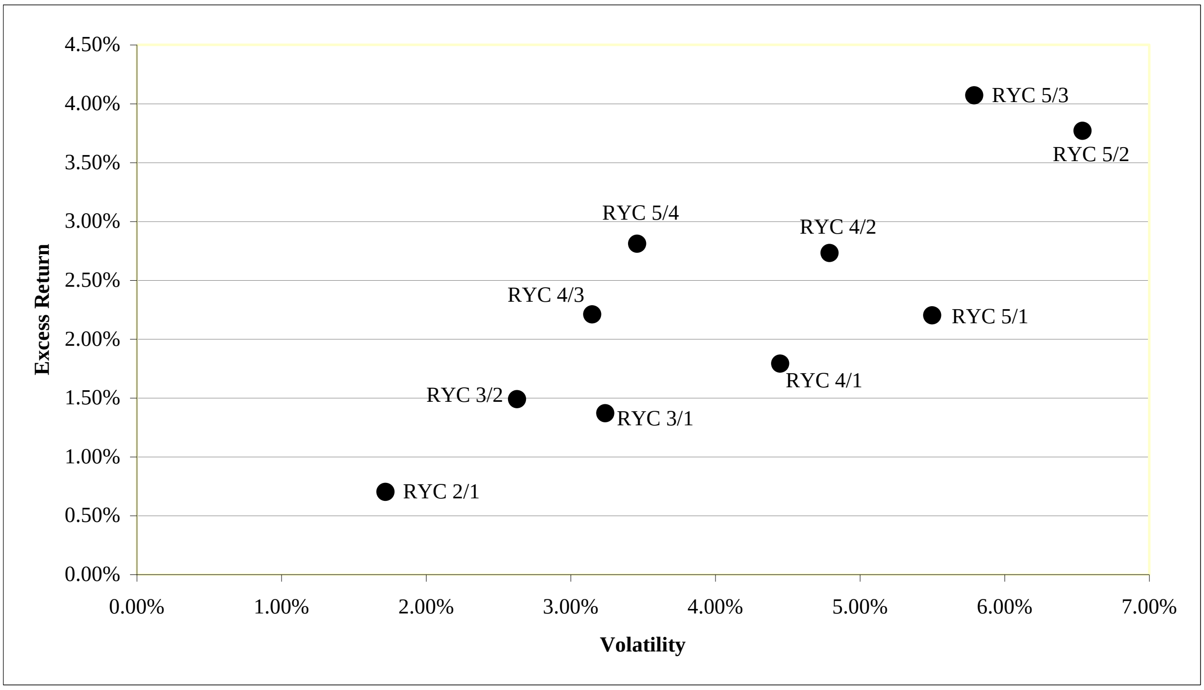
4.4.3 Excess Return - Volatility Profile

The relation between risk and return of unconditional riding strategies is visualized in figure 16. As a risk measure, the standard deviation of the strategy's returns is used. Figure 16 serves as a graphical representation of a performance measure called Sharpe ratio, which is the most well-known performance indicator elaborated by Sharpe (1966). It is the ratio of the excess return of an asset to its standard deviation and is defined as:

$$S_i = \frac{\mathrm{E}(R_i) - r_f}{\sigma_i}, \tag{65}$$

where S_i is the Sharpe ratio and r_f is the risk-free rate of return. The best investment opportunity is that with the highest Sharpe ratio. As can be seen from figure 16, within particular maturities, strategies involving longer holding periods result in better risk-return tradeoff. For example, RYC 3/2 dominates RYC 3/1, as it is both less volatile and provides a higher excess return. The same holds for RYC 4/3, which is a dominant strategy compared to RYC 4/1 and RYC 5/3 which dominates 5/2. Thus, on a risk-adjusted basis, it is generally better to choose a longer holding period within a particular maturity.

Figure 16: Excess Return -Volatility Profile of the Unconditional RYC Strategies



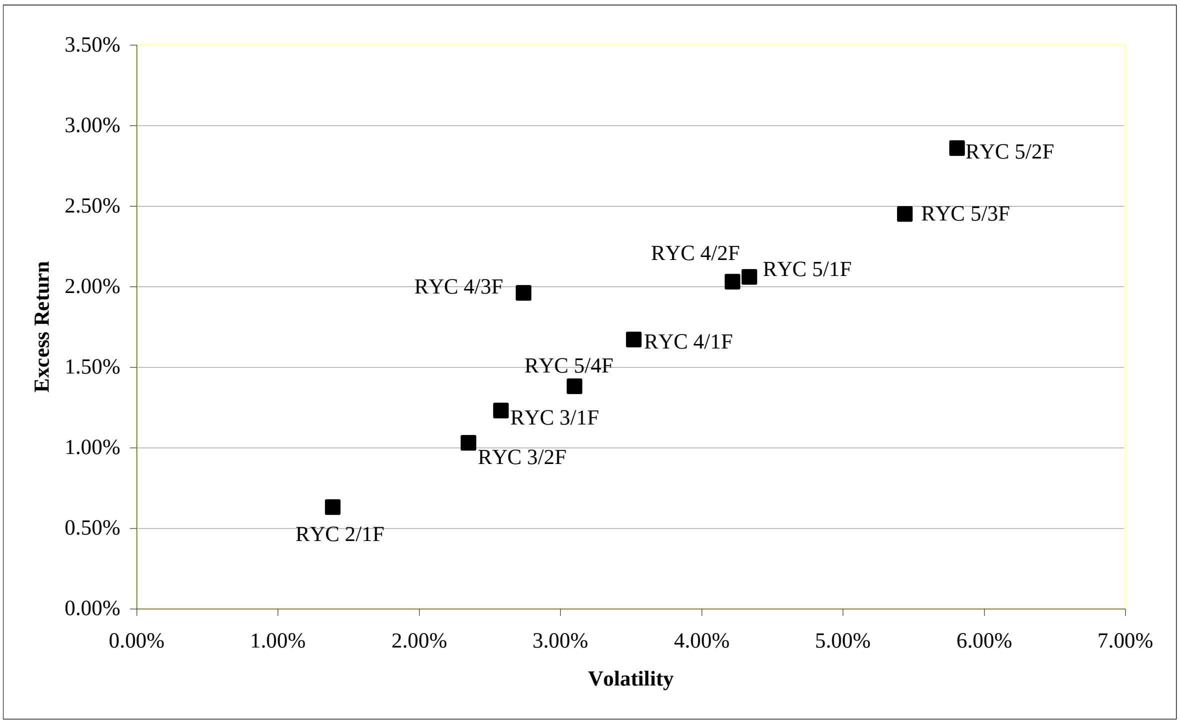
The summary statistics of riding strategies with filter presented in table 24 indicate that the excess returns are more modest than those of the unconditional strategies. The greatest excess return arises from riding a five-year bond for two years and constitutes 2.86 percent compared to the highest return of the unconditional strategies of 4.07 percent. The standard deviations, however, are also lower. Similarly to the unconditional riding, the excess returns of filtered strategies rise along with the maturity of the respective instrument. Regarding the holding period, the returns and standard deviations tend to rise when moving from the shortest holding period to the next one, but then decrease again. For example, for the riding strategy using a four-year bond with a holding period of one, two and three years, the excess returns constitute 1.67, 2.03 and 1.96 percent, respectively. The success rates, though, tend to fall with the longer holding horizon. Figure 17 depicts the excess return/standard deviation profile of conditional riding strategies.

Table 24: Summary Statistics of the RYC Strategy with Filter

Strategy	Mean	Std. Dev.	Minimum	Maximum	Success Rate	Frequency
RYC 2/1F	0.63%	1.39%	-2.94%	5.72%	58.78%	81.5%
RYC 3/1F	1.23%	2.58%	-5.40%	10.78%	53.77%	71.0%
RYC 3/2F	1.03%	2.35%	-6.11%	7.28%	51.51%	71.0%
RYC 4/1F	1.67%	3.52%	-7.68%	13.26%	51.95%	70.6%
RYC 4/2F	2.03%	4.22%	-9.34%	11.98%	68.46%	70.6%
RYC 4/3F	1.96%	2.74%	-6.52%	8.82%	58.81%	70.6%
RYC 5/1F	2.06%	4.34%	-9.84%	15.31%	52.44%	70.1%
RYC 5/2F	2.86%	5.81%	-12.74%	15.75%	51.26%	70.1%
RYC 5/3F	2.45%	5.44%	-13.10%	17.19%	50.26%	70.1%
RYC 5/4F	1.38%	3.10%	-7.63%	10.68%	49.73%	70.1%

Note: The first number next to the name of the strategy denotes the maturity of a fixed income investment whereas the second number stands for the holding period; an “F” means that the strategy was based on a filter rule.

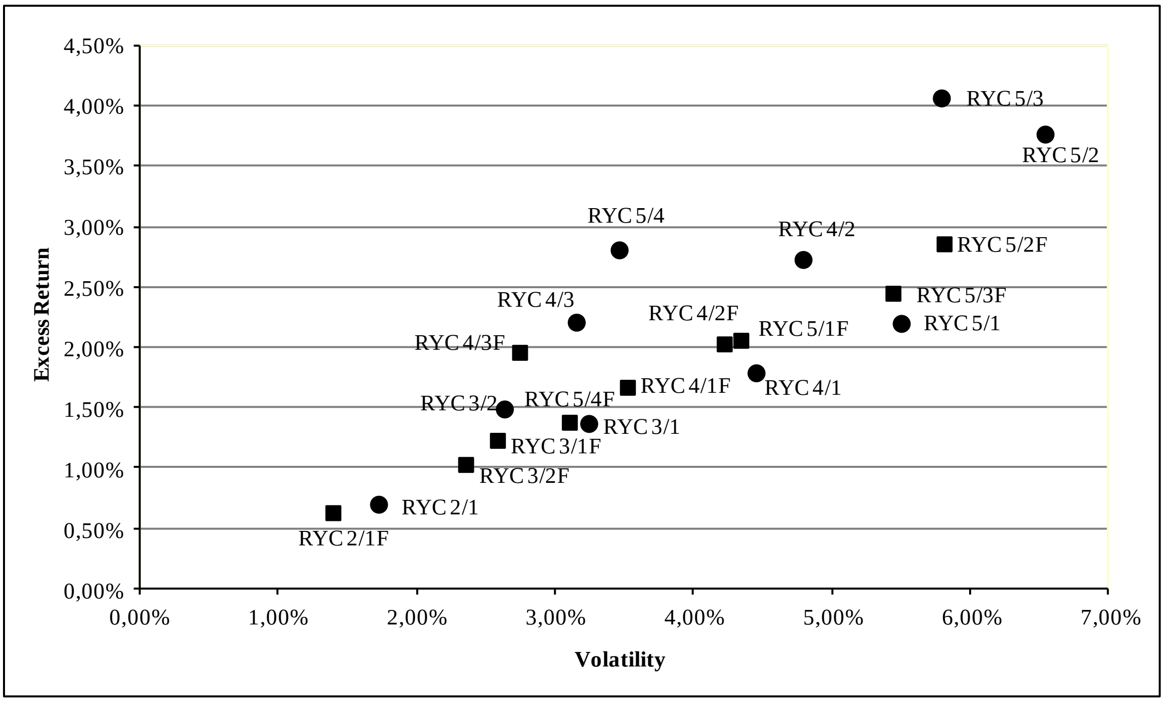
Figure 17: Excess Return -Volatility Profile of the Conditional RYC Strategies



It is evident from figure 17 that riding a four-year bond for three years offers the best opportunities regarding the risk-adjusted return. For the rest of the conditional strategies, there is no such a clear pattern as in case of unconditional riding. However, whereas unconditional riding strategies with longer holding periods generally did better than those with longer ones, the situation seems to be the opposite for riding with filter. For example, RYC 5/1 F, although not dominating strategy RYC 5/3 F, provides a slightly lower excess return, but a much lower volatility. The same is true for RYC 4/1 F compared to RYC 4/2 F. Thus, conditional strategies tend to perform better when short holding periods are applied. In addition, significance tests were conducted for all riding strategies.

Figure 18 depicts the excess return/standard deviation profiles of the conditional as well as unconditional RYC strategies. Comparing to the unconditional strategies, conditional strategies with long maturities and holding periods tend to provide lower standard deviations, but also much lower excess returns. This particularly applies to RYC 5/4 and RYC 5/3. In contrast, for the shortest holding horizon of one year, the conditional strategies generally provide a slightly lower excess return, but a substantially lower standard deviation, independently of the considered maturity. Looking at figure 18, this becomes especially obvious for RYC 5/1 and RYC 4/1.

Figure 18: Excess Return -Volatility Profile of the RYC Strategies



The next section evaluates the risk-return profile of the conditional and unconditional RDYC and RYC strategies with the help of several commonly applied performance measures.

4.5 Performance Evaluation

4.5.1 Performance Measures

In order to evaluate the success of the RDYC and RYC strategies, it is essential to consider not only their rate of return, but also risk associated with the implementation of the strategies. To achieve this, various risk-adjusted performance measures like the above mentioned Sharpe ratio can be applied. The following three sub-sections describe the traditional performance evaluation techniques as well as recently developed value-at-risk- and lower partial moment-based measures.

4.5.1.1 Traditional Performance Measures

The Sharpe ratio is very easy to apply and is not benchmark-related, i.e. it is possible to compare two investment alternatives without a benchmark specification. However, the Sharpe ratio has been criticized recently because of its reliance on variance as risk measure. Traditional performance measures such as the Sharpe ratio are based on the mean-variance portfolio theory, according to which investors' preferences can be described using mean and standard deviation of portfolio returns. In turn, according to the expected utility theory, investors make their decisions by maximizing their expected utility. The $\mu-\sigma$ analysis is then an appropriate framework if it is consistent with the expected utility theory, which has a solid theoretical foundation and represents a standard approach in economics to making decisions under uncertainty. If this is the case, then variance represents an appropriate risk measure. However, the mean-variance framework is only consistent with the expected utility theory in the case that the investors' utility function can be described as quadratic or asset returns are normally distributed.[97] In the case that these conditions are not met, the Sharpe Ratio could make false indications with respect to the asset's performance.[98]

[97] Or follow a more general class of elliptical distributions (see Ingersoll (1987)).

[98] This point was indicated by many authors, especially in the hedge funds literature. See, for example, Brooks and Kat (2002), Kao (2002), Gregoriou and Gueyie (2003). However, not all researchers share this view. Eling and Schuhmacher (2007) evaluate hedge funds performance using traditional as well as modern performance measures and report that the choice of the measure did not matter for funds ranking. Eling (2008) confirmed these findings for the ranking of mutual funds. Several authors stress the problematic interpretation of negative Sharpe ratios: Israelsen (2003) shows that in case of negative excess returns Sharpe Ratio gives higher ranking to funds with both smaller (more negative) excess returns and higher standard deviation than other funds. Similar results regarding funds ranking with negative Sharpe ratios were obtained by Scholz (2007).

For this reason, several alternative approaches to measure risk have been proposed.[99] Among them, the so called downside performance measures have drawn special attention of the researchers. In these measures, variance as a risk measure has been replaced by downside risk measures such as the Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR) as well as the Lower Partial Moment (LPM).

4.5.1.2 Value at Risk-Based Performance Measures

Value at Risk (VaR) represents the maximum loss of an investment that will not be exceeded with the probability 1 – *g* during a certain period.[100] The maximum loss can be stated either as some particular amount or as a rate of return. Being a convenient way to summarize the risk of a financial institution in one single figure, VaR has become a standard risk measurement tool for financial institutions.

Crucial for its popularity was the introduction of the amendment to the original Basel Capital Accord of 1988 by the Basel Committee for Banking Supervision. While the original Basel Accord contained capital requirements to cover credit risk, the 1996 amendment adopted the VaR approach for evaluating market risks. Regulatory requirements defined the VaR as the maximum loss encountered over the period of ten days at the 99 percent confidence level. If *F* is a continuous distribution function of the portfolio value, the value at risk is computed as the *g*-th quantile of the inverse of the distribution function *F*:

$$VaR_g = -F^{-1}(g). \qquad (66)$$

There are three main approaches to the VaR calculation: the parametric approach, the historical simulation approach and the Monte Carlo Simulation approach. The framework of the parametric approach assumes the normal distribution of returns and requires the estimation of the mean and the standard deviation parameter from the empirical returns distribution. In this case, the VaR is computed using the formula $VaR_g = -(\overline{R}_i + z_g \cdot \sigma_i)$, where z_g is the *g*-th quan-

---

[99] Some modifications of the Sharpe Ratio include Risk-adjusted Performance (RAP) proposed by Modigliani and Modigliani (1997), Market Risk-Adjusted Performance (MRAP) found in Wilkens and Scholz (1999) and Correlation-Adjusted Performance (CAP) suggested by Muralidhar (2000).

[100] See Jorion (1997), pp. 85-91.

tile of the standard normal distribution.[101] However, if the returns distribution deviates from normality and exhibits negative skewness and fat tails, such an estimation may yield misleading results.[102] Addressing this issue, the Cornell-Fisher approximation was proposed to adjust the VaR estimate and take into consideration the third and the fourth moment of a distribution, i.e. skewness and kurtosis. The latter is then referred to as modified value at risk (MVaR) and is computed as:

$$\mathrm{MVaR} = -\left(\overline{R}_i + \sigma_i \cdot \left[z_g + \frac{1}{6}\left(z_g^2 - 1\right)S_i + \frac{1}{24}\left(z_g^3 - 3z_g\right)K - \frac{1}{36}\left(2 \cdot z_g^3 - 5 \cdot z_g\right) \cdot S_i^2 \right]\right) \qquad (67)$$

where S_i denotes skewness and K_i stands for excess kurtosis of the asset *i*. It is worth noting that the above equation is just an approximation around the normal distribution and is therefore suitable only in case of moderate skewness and excess kurtosis.[103] If deviations from normality are large, the Cornell-Fisher approximation should not be applied.

The historical simulation approach uses the actual empirical returns distribution over a sufficiently long time horizon to estimate possible losses. Finally, the Monte Carlo simulation uses the empirical distribution in order to run a large number of simulations and estimate VaR and, therefore, represents the most computationally intensive method.[104] The main point of criticism concerning the VaR is that it does not satisfy all necessary criteria to be considered an acceptable risk measure. According to Artzner et al. (1999), a risk measure *p* that belongs to the class of *coherent risk measures* must satisfy four conditions for all random variables *x* and *y*:

1) Subadditivity: $p(x+y) \le p(x) + p(y)$;
2) Homogeneity: $p(\lambda \cdot x) = \lambda \cdot p(x)$ for all positive real numbers λ ;
3) Monotonicity: $p(x) \ge p(y)$ if $x \ge y$;
4) Transitional invariance: $p(x + n \cdot r_f) = p(x) - n$.

[101] See Crouhy/Galai/Mark (2009), pp. 154-155.
[102] Favre and Galeano (2002) demonstrate that using VaR to evaluate the performance of hedge funds leads to serious underestimation of risk.
[103] See Lhabitant (2004), p. 299.
[104] See Fabozzi/Mann (2010), pp. 380-381.

Whereas VaR fulfills the last three conditions, it does not possess the subadditivity property[105] and therefore cannot be considered a coherent risk measure. The lack of subadditivity means that the risk of a combination of two portfolios can be greater than the sum of risks of individual portfolios, i.e. $VaR(P_1+P_2)\leq VaR(P_1)+VaR(P_2)$. Only in case of normal distribution VaR can be called a coherent risk measure.[106] This is an important point of criticism, stressed by many authors, especially in the contest of adopting VaR as a measure to calculate the regulatory capital requirements. The fact that VaR is not subadditive means it would punish portfolio diversification, and lead to the choice of a less diversified portfolio, as more assets included into the portfolio would also mean higher risk.[107] Another argument is that VaR would motivate splitting up a company into several parts, in order to diminish risks.[108] An additional point of criticism regarding VaR is that it is not able to capture the expected loss beyond the specified confidence level.

As a response to the above mentioned drawbacks, another measure initially introduced as Expected Shortfall in Acerbi (2001) and later mentioned in Rockafellar and Uryasev (2002) as Conditional Value at Risk (CVaR) has been proposed as an alternative. It represents the expected loss exceeding the VaR and is defined as follows:

$$\mathrm{CVaR}_g = \mathrm{E}\left(R \middle| R \leq \mathrm{VaR}_g\right) \tag{68}$$

As opposed to VaR, the CVaR[109] fulfills all requirements of a coherent risk measure including the sub-additivity.[110] In addition, it allows taking into consideration extreme events measured by the lower tail of the distribution, which are not captured by VaR.[111] In case of normally distributed returns, both measures lead to the same result. Based on VaR, MVaR and CVaR, different performance indicators have been created. They include the excess return on value at risk (ERVaR) introduced by Dowd (2000), the modified Sharpe ratio (MSR) derived in Gre-

---

[105]Artzner et al. (1999), pp. 216-218, were the first to demonstrate that this property is not fulfilled by VaR. Later, some numerical examples of VaR lacking subadditivity were given in the literature; see, among others Tasche (2002), p. 1522, Frey and McNeil (2002), pp. 1321-1322.

[106] Embrechts et al. (2002) show that this result is also valid for a general class of elliptical distributions.

[107] Danielsson (2002), p. 1289, provides an example illustrating that using VaR to choose between a diversified and non-diversified portfolio would lead to the choice in favour of the latter.

[108] See Szegö (2002), p. 1260.

[109] CVaR corresponds to the Lower Partial moment of second degree, which will be introduced in the next section.

[110] See Artzner et al. (1997), pp. 68-69.

[111] See Hull (2007), pp. 198-199.

goriou and Gueyie (2003) and the conditional Sharpe ratio (CSR) of Agrawal and Naik (2004):

$$\mathrm{ERVaR}_g^i = \frac{\overline{R}_i - r_f}{\mathrm{VaR}_g} \tag{69}$$

$$\mathrm{MSR}_g^i = \frac{\overline{R}_i - r_f}{\mathrm{MVaR}_g} \tag{70}$$

$$\mathrm{CSR}_g^i = \frac{\overline{R}_i - r_f}{\mathrm{CVaR}_g} \tag{71}$$

4.5.1.3 Lower Partial Moment-Based Performance Measures

Lower partial moments belong to the so called downside risk measures and were presented in their general form in Bawa (1975) and Fishburn (1977). Recent popularity of lower partial moments as a risk measure is connected with their favorable theoretical properties: the main advantage of this measure is that it is consistent with the expected utility theory and is applicable to a broad class of return distributions. Moreover, LPMs are compatible with the principle of the stochastic dominance,[112] which will be introduced in section 4.3.2.4. Lower partial moments focus only on the left tail of the return distribution and do not treat positive deviations as risky, which is the case if volatility is used. Instead, it gives a possibility to take into account only negative deviations from some target rate of return, τ. The lower partial moment is generally defined as:

$$\mathrm{LPM}_q(\tau) = \left[E(\tau - R_i)^q \middle| \tau > R_i\right]^{\frac{1}{q}}, \tag{72}$$

where $\mathrm{LPM}_q(\tau)$ denotes the lower partial moment of order q with the minimum acceptable return τ .Typically, some minimum acceptable return, such as the risk-free rate of return, zero rate of return or the mean return serve as the target rate. Thus, LPM is a risk measure that is especially appealing to investors whose main goal is to achieve some target rate of return, such as pension funds. For discrete random variables, the lower partial moment for T observations is computed in the following way:

[112] This was shown in Bawa (1975, 1978), Fischburn (1977), Nawrocki (1991, 1999).

$$\mathrm{LPM}_q(\tau) = \frac{1}{T}\sum_{t=1}^{T}\max\left(\tau - R_i;0\right)^q. \tag{73}$$

The order of the lower partial moment measures the extent to which returns below the target rate of return are penalized and, thus, reflects the investor's perception of risk. Higher orders of the LPM put higher weight on the negative deviations from the target rate of return and, thus, reflect the investor's higher sensitivity to such deviations. Whereas $0<q<1$ stands for a risk-seeking behavior, $q>1$ describes a risk-averse investor with respect to returns below the target rate.[113] The most well-known are the lower partial moments of order zero, one and two. Based on the LPM of different orders, several performance measures have been developed. The Omega measure was introduced by Shadwick and Keating (2002). It is based on the LPM of order one and computes the excess return as the difference between the mean return and a prespecified threshold τ . The Omega is defined in the following way:

$$\mathrm{Omega}_i = \frac{\overline{R}_i - \tau}{\mathrm{LPM}_1(\tau)}. \tag{74}$$

Sortino and Price (1994) developed a performance measure based on the LPM of the second order. Later, the measure became known as the Sortino ratio and represents a downside equivalent to the Sharpe ratio. It measures excess return per unit of downside risk and is given by:

$$\text{Sortino Ratio}_i = \frac{\overline{R}_i - \tau}{\sqrt{\mathrm{LPM}_2(\tau)}}. \tag{75}$$

Despite the fact that the Sortino ratio has been a subject of criticism by its author himself,[114] it has become a popular way to measure risks, especially in the hedge funds literature. Kappa of order q was introduced by Kaplan and Knowles (2004) as a general risk measure based on the LPM, which is defined as follows:

[113] This was shown by Fishburn (1977).
[114] See Sortino/Kordonsky/Forsey (2006), p. 5.

$$\text{Kappa}_q = \frac{\overline{R}_i - \tau}{\sqrt[q]{\text{LPM}_q(\tau)}}. \tag{76}$$

The Omega and the Sortino ratio can be easily derived from equation (76). The Kappa of the third order corresponds to the LPM of the third order and is given by:

$$\text{Kappa}_3 = \frac{\overline{R}_i - \tau}{\sqrt[3]{\text{LPM}_3(\tau)}}. \tag{77}$$

Kaplan and Knowles (2004) analyze the performance of hedge funds employing the Kappa measure of order one, two and three. They demonstrate that the choice of the order had significant influence on the hedge funds' ranking. An additional performance measure derived by Sortino, van der Meer and Plantiga (1999) is referred to as the upside potential ratio (UPR). Its origins go back to the study of Fishburn (1977) who stated that individuals strongly dislike returns lying below the target return and show risk-averse behavior in this case. However, the more the returns exceed the threshold, the more they like them and therefore are risk-neutral in this case. The ratio measures upside potential relative to downside risk and uses an upper partial moment (UPM) in the numerator, which measures positive deviations from the target rate of return:

$$\text{UPR} = \frac{\text{UPM}_1(\tau)}{\sqrt{\text{LPM}_2(\tau)}}. \tag{78}$$

where UPM_1 represents an upper partial moment of the first degree. The UPM of order q relative to a target rate of return τ is given as:

$$\text{UPM}_q(\tau) = \frac{1}{T}\sum_{t=1}^{T}\max\left(\text{R}_\text{i} - \tau;0\right)^q. \tag{79}$$

The UPM maximizes the expected returns above the threshold and represents potential for success.[115] In the denominator of the UPR is the LPM of the second order, which represents

[115] See Sortino/van der Meer/Plantiga (1999), p. 54.

the downside risk. It must be noted that UPM, as opposed to the LPM, does not have any decision theory-based foundation.

4.5.1.4 Stochastic Dominance Criterion

Another measure that can be used to rank two alternative investment opportunities is the stochastic dominance criterion. It is based on the comparison of the cumulative distribution functions (cdf) associated with two strategies. The cdf, denoted as $F(x)$, is defined by:[116]

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx, \tag{80}$$

where $f(x)$ is the density function. The investors' utility function is represented by u_1 whose first derivative is non-negative ($u' \geq 0$). This implied that investors prefer more to less. If two investments F and G are to be ranked whose cumulative probability functions are $F(x)$ and $G(x)$, then F dominates G by the *first order stochastic dominance* if the following condition is fulfilled:[117]

$$F(x) \leq G(x) \tag{81}$$

for all x and there is at least one x for which inequality (81) holds strictly. Graphically, if the first order stochastic dominance exists, the cumulative probability distributions of two investment opportunities should not cross, but may have a tangency point. Then F dominates G, and its cdf lies below that of investment opportunity G. In this case, compared to G, F assigns a lower probability to obtain a value lower than x.[118] If there exists a crossover point and the first order stochastic dominance does not apply, the *second order stochastic dominance* may be considered. Here, an assumption is made that investors prefer more to less and are risk-averse. Their utility function is represented by u_2 ($u' > 0; u'' < 0$). According to this criterion, F dominates G by the second order stochastic dominance if the following holds:[119]

[116] See Levy (2006), p. 53.
[117] See Ingersoll (1987), p. 71.
[118] See Levy (2006), pp. 59-61; p. 64.
[119] See Ingersoll (1987), p. 137.

$$\int_{-\infty}^{x} F(x)dx \leq \int_{-\infty}^{x} G(x)dx \tag{82}$$

for all *x* and there is at least one *x* for which the inequality (82) holds strictly. If *F* dominates *G* by the second order stochastic dominance, the area under the cumulative probability function of *F* should be smaller than the area under the cumulative distribution function of *G*. The second order stochastic dominance is less restrictive than the first order stochastic dominance. The existence of the first order stochastic dominance implies that there is also a stochastic dominance of a higher order. Moreover, if the second order stochastic dominance is not applicable, it is possible to derive the conditions for stochastic dominance of higher orders taking the integral of inequality (82). For example, the condition for the third order stochastic dominance would be given by:

$$\int_{-\infty}^{x} \int_{-\infty}^{y} F(x)dydy \leq \int_{-\infty}^{x} \int_{-\infty}^{y} G(x)dydx \tag{83}$$

for all *x* and there is at least one *x* for which the inequality (83) holds strictly. Here, an assumption is made that investors prefer more to less, are risk-averse and exhibit decreasing absolute risk aversion, i.e. as wealth increases, they will hold more money in risky assets. Their utility function is represented by u_3 ($u' > 0; u'' < 0; u''' > 0$). The concept of the stochastic dominance is especially appealing because it does not require any assumption about the distribution function of returns and only some general assumptions regarding the investors' utility function.

4.5.2 Rolling Down the Yield Curve Strategy

In order to examine the distributional characteristics of the returns to the RDYC strategies, the Jarque-Berra test will be applied. The actual performance of the rolling strategies will be then measured with the help of the Sharpe ratio, excess return on VaR, CSR, Omega, Sortino ratio, UPR and the stochastic dominance criterion. Table 25 provides the distributional characteristics of unconditional and conditional strategies including the mean, the standard deviation, the third and the fourth moments of the distribution, skeweness and kurtosis.

Table 25: Descriptive Statistics of the RDYC Strategy

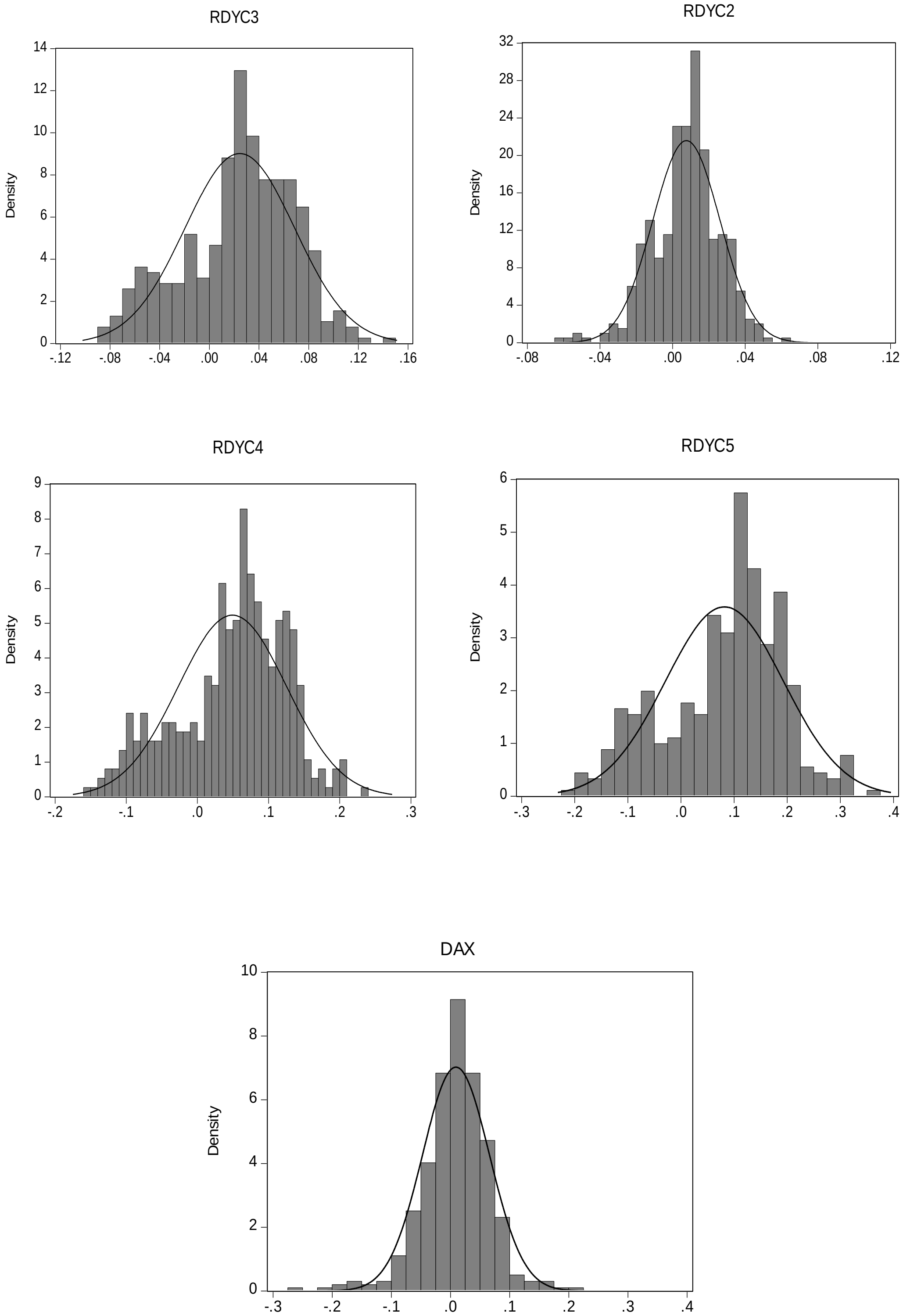
	Mean Excess Return	Std. Dev.	Skewness	Kurtosis	Jarque-Berra	p-value
			Unconditional			
RDYC 2	0.76%	1.84%	-0.466	3.869	26.945	0.0000
RDYC 3	2.451%	4.42%	-0.394	2.670	11.739	0.0028
RDYC 4	4.926%	7.63%	-0.493	2.742	16.120	0.0030
RDYC 5	8.179%	11.13%	-0.420	2.705	11.943	0.0026
DAX	0.420%	5.68%	-0.481	5.483	117.600	0.0000
			Conditional			
RDYC 2F	0.69%	1.47%	0.2029	3.34	4.6	0.0990
RDYC 3F	1.91%	3.99%	-0.0822	3.11	0.64	0.7259
RDYC 4F	3.37%	6.83%	-0.0377	2.00	0.088	0.9566
RDYC 5F	4.93%	10.04%	0.1370	2.90	1.31	0.5220
DAX	0.420%	5.68%	-0.4810	5.483	117.600	0.0000

The number next to the name of the strategy denotes the number of years of the long-term investment; an "F" means that the strategy was based on a filter rule.

As table 25 shows, the excess returns of all unconditional strategies are characterized by negative skewness; one strategy, RDYC 2, also exhibits excess kurtosis. Two of the conditional strategies, RDYC 3 and RDYC 4, are slightly negatively skewed whereas RDYC 2 and RDYC 3 have excess kurtosis. As for the excess returns of the DAX, they are negatively skewed and have excess kurtosis. This can be also conveniently observed from figure 19 that provides a histogram of the excess returns of unconditional RDYC strategies as well as the DAX in comparison with the normal distribution. For example, in case of the DAX, figure 19 shows that the probability of the extreme events to occur is higher than for the normal distribution, i.e. there is a "fat tails" problem. The results of the Jarque-Berra test indicate that the returns of all unconditional strategies and the DAX do not follow a normal distribution. The *p*-values of the JB statistics are sufficiently low, so that one can reject the null hypothesis of a normal distribution at the one percent level. In case of conditional strategies, the null hypothesis of a normal distribution cannot be rejected for RDYC 3, RDYC 4 and RDYC 5, as the *p*-value is too high. For RDYC 2, it can only be rejected at the ten percent significance level. Thus, the returns of the conditional strategies follow a normal distribution in most of the cases whereas all unconditional strategies deviate from normality.

Having investigated the distributional characteristics of the returns, we turn to the actual performance of the rolling strategies. It will be measured with the help of the Sharpe ratio, excess return on VaR, CSR, Omega, Sortino ratio and the UPR. Although each of these measures has

Figure 19: Histogram of the RDYC Strategies with Normal Distribution



its own specifics, we use all above mentioned performance indicators. This will allow making a judgment on the basis of several different measures of risk, such as standard deviation, VaR, CVaR and the LPMs. For performance indicators based on lower partial moments, the value of τ equal to zero is taken. In case of VaR-based measures, the value at risk is calculated with the parametric approach at the five percent significance level. In addition, it is interesting to compare the rankings resulting from different performance measures. The performance results of the conditional as well as unconditional strategies are presented in table 26.

Table 26: Performance of the RDYC Strategy

Strategy	Sharpe ratio	Excess return on VaR	Conditional Sharpe ratio	Omega	Sortino ratio	Upside potential ratio
Unconditional						
RDYC 2	0.4114	0.2144	0.1508	1.79	0.793	1.15
RDYC 3	0.5541	0.3120	0.2910	2.62	1.126	1.55
RDYC 4	0.6458	0.3834	0.3483	3.33	1.382	1.80
RDYC 5	0.7350	0.4610	0.4276	4.23	1.753	2.17
DAX	0.0700	0.0339	0.0249	0.24	0.111	0.66
Conditional						
RDYC 2F	0.4052	0.0878	0.0754	2.64	1.031	1.42
RDYC 3F	0.4774	0.1976	0.1651	2.90	1.000	1.34
RDYC 4F	0.4933	0.2331	0.2113	3.12	1.069	1.41
RDYC 5F	0.4912	0.2430	0.2313	3.11	1.128	1.49
DAX	0.0700	0.0339	0.0249	0.24	0.111	0.66

The number next to the name of the strategy denotes the number of years of the long-term investment; an "F" means that the strategy was based on a filter rule.

For unconditional strategies, all considered performance measures indicate better performance when moving to a longer-maturity instrument. For example, the Sharpe ratio rises from 0.4114 to 0.5541 from the two-year strategy to a three-year strategy and constitutes 0.7350 for a five-year strategy. That means, the highest performance among the unconditional strategies is achieved when rolling down with the five-year bond. The two-year strategy performs the worst in accordance with all performance indicators. It is worth noting, that all considered unconditional strategies outperformed the DAX, whose Sharpe ratio only constitutes 0.07.

This outperformance of the rolling strategies is also indicated by all other measures. For strategies where the filter rule was applied, most of the measures rise with the maturity of the applied instrument. The only exception is the three-year strategy, which, according to the Sortino ratio and the UPR, has done worse than the two-year strategy. Although the ratios rise with a longer maturity, this rise is not as substantial as in the case of unconditional strategies. Namely, the Sharpe ratio increases from 0.4052 for a two-year strategy to only 0.4912 for the five-year strategy, compared to the increase from 0.4114 to 0.7350 for the unconditional strategies. Similar results can be reported for all other performance indicators, which do not rise that sharply as for strategies without filter. It is also of interest to compare the unconditional strategies with the conditional ones in terms of the achieved performance. Table 27 provides the ranking of all RDYC strategies in accordance with the considered performance indicators.

Table 27: Ranking of the RDYC Strategies

Strategy	Sharpe ratio	Excess return on VaR	Conditio-nal Sharpe ratio	Omega	Sortino ratio	Upside potential ratio
RDYC 5	1	1	1	1	1	1
RDYC 4	2	2	2	2	2	2
RDYC 3	3	3	3	7	4	3
RDYC 5F	4	4	4	4	3	4
RDYC 4F	5	5	5	3	5	6
RDYC 3F	6	6	6	5	7	7
RDYC 2	7	8	8	8	8	8
RDYC 2F	8	7	7	6	6	5
DAX	9	9	9	9	9	9

Note: The number next to the name of the strategy denotes the number of years of the long-term investment; an "F" means that the strategy was based on a filter rule.

The best performing strategy is the five-year strategy without a filter rule, which is indicated by all performance measures in table 27. The five-year strategy is followed by the four and three-year unconditional strategies, although the latter is on the third place according to all measures except of the Sortino ratio and the Omega, which assign the fourth and the seventh place, respectively. In contrast, the DAX possesses the worst performance characteristics according to all measures.

Table 27 confirms the previous finding that unconditional strategies perform better than the conditional ones in most cases. According to the Sharpe ratio, strategies without the filter outperform the conditional strategies for all considered maturities. The four- and the five-year unconditional strategies outperform the conditional ones according to all measures. The three-year unconditional strategy exhibits superior performance over the five-year conditional strategy according to all but one measure. Finally, in case of the two-year conditional strategy, it outperforms the unconditional two-year strategy according to the Omega, the Sortino ratio and the UPR, but not the ERVaR or CVaR. Thus, the filter rule proved to be ineffective for all maturities except of, in some cases, the shortest maturity of two years. Thus, one is better off performing the rides all the time, independent of the term structure at the moment of the strategy's initiation. Relying on a positive term structure does not enhance performance, as the fact that the term structure is positive at time *t* does not seem to be capable of predicting the future term structure of interest rates.

As table 27 shows, there exist certain differences among various performance indicators. Whereas the Sharpe ratio, the ERVaR and the CSR almost provide identical rankings, there are some stronger differences between other measures. In this case, the rank correlation coefficient can be calculated to measure divergence between alternative performance measures. The Spearman's rank correlation coefficient can be used for this purpose. It is computed in the following way:[120]

$$\text{Rank correlation} = 1 - 6 \cdot \left(\frac{\sum_{i=1}^{n} d_i^2}{n \cdot \left(n^2 - 1\right)} \right), \tag{84}$$

where d_i is the distance between two rankings of the strategy i and n is the number of the strategies. Table 28 contains the values of the rank correlation coefficients, which are estimated pair-wise for all measures. There exists a high rank correlation between the Shape ratio and other performance measures, as rankings based on the Sharpe ratio are at least 78.3 percent correlated with the rankings according to other measures, the smallest correlation being between the Sharpe ratio and the Omega. With the rest of all measures, the correlation is at least 90 percent. The correlation between other measures is also high, as the correlation coefficients lie in the range between 75 percent and 100 percent, the lowest being between the Omega and the UPR. Rankings according to the ERVaR and the CSR are identical. Thus, the choice of the performance measure does not influence the assigned ranking significantly.

[120] See Spearman (1904), pp. 72-101.

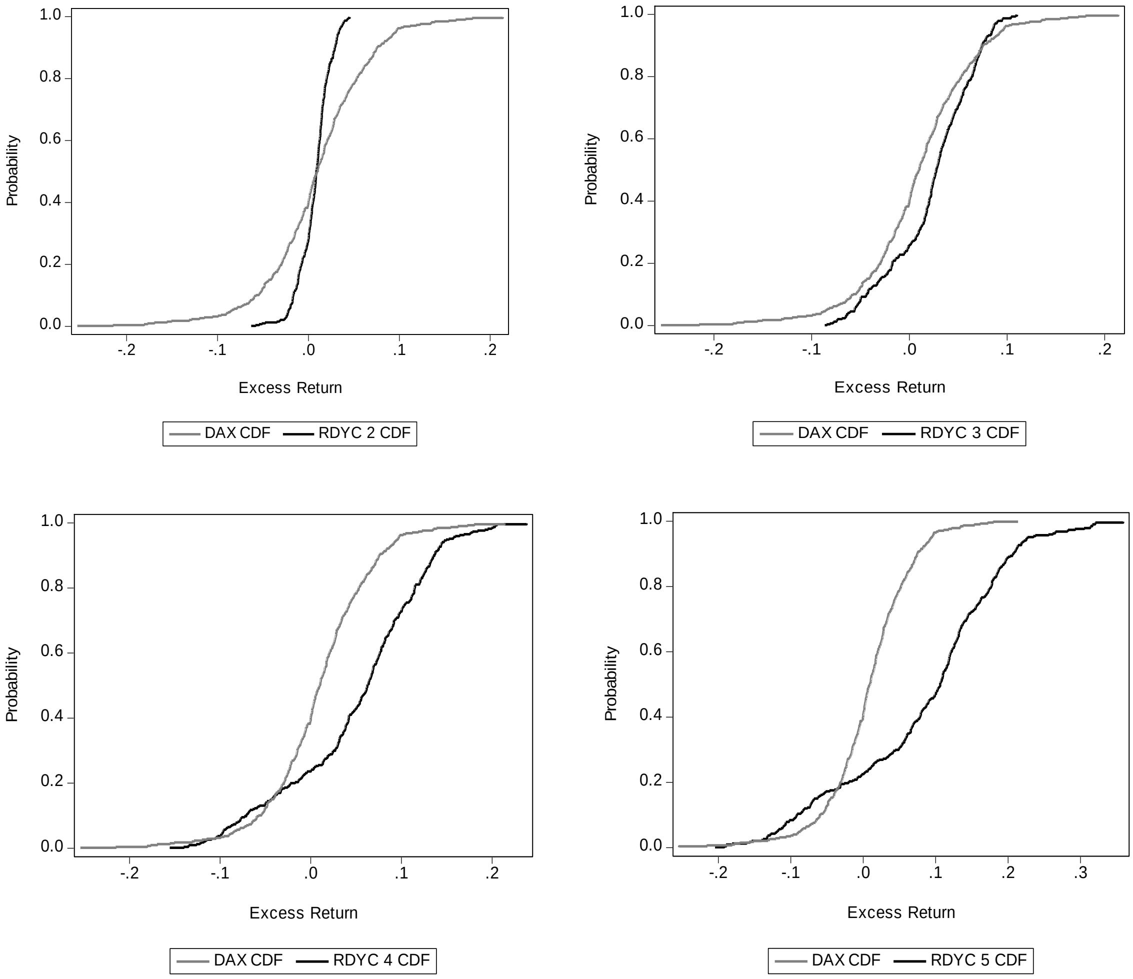
Table 28: Rank Correlation of Performance Measures (RDYC)

Strategy	Sharpe ratio	Excess return on VaR	Condi tional SR	Omega	Sortino ratio
UPR	0.900	0.950	0.950	0.750	0.967
Sortino Ratio	0.933	0.967	0.967	0.850	
Omega	0.783	0.817	0.817		
Conditional SR	0.983	1.000			
Excess return on VaR	0.983				

Figure 20 plots the cumulative probability distribution of returns from four RDYC unconditional strategies and the DAX returns. As can be seen from this figure, there exists a crossover point of two distributions in every case. Consequently, the concept of the first order stochastic dominance cannot be applied. However, for RDYC 3, 4 and 5 the area below the cumulative distribution function is smaller than that of the DAX. Thus, in most cases rolling strategies dominate the DAX by the second order stochastic dominance.

The RDYC strategy outperforms the capital market investment strategy. Moreover, most rolling strategies dominate the DAX by the second order stochastic dominance. Yet, the RDYC strategy is not risk-free. The main source of risk is that, contrary to the expectations, the term structure does not stay stable and interest rates rise. If this happens, the strategy could suffer losses, as refinancing could become more expensive. One implication of that would be trying to hedge this risk using appropriate financial instruments. From a theoretical point of view, however, investors only get a return above the risk-free return if the investment is risky. Applying some hedging instruments would, thus, lead to only a risk-free rate of return on the RDYC strategy. Consequently, it is not correct to call the RDYC a pure arbitrage strategy. It is rather a “risk arbitrage” strategy, as it provides a significantly higher risk-return tradeoff than the market level.

Figure 20: Cumulative Probability Distribution of RDYC strategies



4.5.3 Riding the Yield Curve Strategy

In case of RYC strategies, the distributional characteristics of the excess return are considered. They are reported in table 29 for unconditional strategies and in table 30 for conditional rides.

Table 29: Descriptive Statistics of the RYC Strategy (Uncondtional)

Strategy	Mean Excess Return	Std. Dev.	Skewness	Kurtosis	Jarque-Berra	*p*-value
RYC 2/1	0.695%	1.72%	-0.322	3.541	12.081	0.0024
RYC 3/1	1.37%	3.24%	-0.300	2.980	5.986	0.0501
RYC 3/2	1.49%	2.63%	-0.475	2.574	16.870	0.0002
RYC 4/1	1.79%	4.45%	-0.276	2.820	5.740	0.0567
RYC 4/2	2.73%	4.79%	-0.488	2.444	20.93	0.0000
RYC 4/3	2.21%	3.15%	-0.627	3.016	25.315	0.0000
RYC 5/1	2.20%	5.50%	-0.312	2.795	7.352	0.0253
RYC 5/2	3.77%	6.54%	-0.483	2.480	19.94	0.0000
RYC 5/3	4.07%	5.79%	-0.552	2.914	19.745	0.0001
RYC 5/4	2.81%	3.46%	-0.426	3.051	11.363	0.0034

Note: The first number next to the name of the strategy denotes the maturity of a fixed income investment, whereas the second number stands for the holding period.

All unconditional riding strategies are skewed to the left; three of them, RYC 2/1, RYC 4/3 and RYC 5/4, also exhibit excess kurtosis. The JB test indicates that the hypothesis about normally distributed returns can be rejected: for the most of the considered strategies at the one percent level; for three strategies the significance levels vary from 2.5 percent to six percent. As for conditional strategies, three of them are negatively skewed and most have excess kurtosis. However, for this type of strategies the normality could sometimes not be rejected. The JB test is not able to reject the null hypothesis of normally distributed return in half of the cases, as the *p*-value of the test statistic is too large. For the rest of the conditional strategies the null could be rejected at least at the five percent level of significance.

Table 30: Descriptive Statistics of the RYC Strategy (Conditional)

Strategy	Mean Excess Return	Std. Dev.	Skewness	Kurtosis	Jarque-Berra	*p*-value
RYC 2/1	0.63%	1.39%	0.280	3.249	6.789	0.0336
RYC 3/1	1.23%	2.58%	0.324	3.251	8.024	0.0181
RYC 3/2	1.03%	2.35%	-0.126	2.792	1.773	0.4121
RYC 4/1	1.67%	3.52%	0.380	3.137	10.116	0.0064
RYC 4/2	2.03%	4.22%	-0.012	2.568	3.105	0.2117
RYC 4/3	1.96%	2.74%	-0.144	2.668	3.139	0.2081
RYC 5/1	2.06%	4.34%	0.362	3.109	9.176	0.0102
RYC 5/2	2.86%	5.81%	0.035	2.602	2.714	0.2574
RYC 5/3	2.45%	5.44%	0.121	3.031	0.9597	0.6189
RYC 5/4	1.38%	3.10%	0.274	3.408	7.286	0.0262

Note: The first number next to the name of the strategy denotes the maturity of a fixed income investment, whereas the second number stands for the holding period.

The performance results of various RYC unconditional strategies are reported in table 31. The performance results for the strategies that were based on the positive slope filter rule are shown in table 32. The excess returns are computed as the difference between the rate of return of the RYC strategy and the rate of return on the corresponding BH strategy. For performance indicators based on LPMs, the value of τ equal to zero is applied. In case of value at risk-based measures, the VaR is calculated with the parametric approach at the five percent significance level. Looking at the Sharpe ratios, it is evident that the performance improves with the holding horizon of the respective instrument. Riding the yield curve with a five-year instrument results in the Sharpe ratio of 0.3995. However, riding the same instrument for four years doubles the Sharpe ratio, which constitutes 0.8113. Although an increase in the Sharpe ratio is especially pronounced in case of the five-year bonds, the substantial improvement occurs also with all other maturities, as one is better off using a longer holding period.

Table 31: Performance of the RYC Strategy (Unconditional)

Strategy	Sharpe ratio	Excess return on VaR	Conditional Sharpe ratio	Omega	Sortino ratio	UPR
RYC 2/1	0.4052	0.2105	0.1554	1.74	0.750	1.18
RYC 3/1	0.4231	0.2219	0.1632	1.79	0.803	1.25
RYC 3/2	0.5651	0.3202	0.2813	2.63	1.168	1.61
RYC 4/1	0.4016	0.2082	0.1573	1.63	0.758	1.22
RYC 4/2	0.5704	0.3241	0.2923	2.63	1.181	1.63
RYC 4/3	0.7013	0.4306	0.3370	3.97	1.546	1.93
RYC 5/1	0.3995	0.2069	0.1665	1.60	0.745	1.21
RYC 5/2	0.5766	0.3289	0.2851	2.71	1.202	1.65
RYC 5/3	0.7038	0.4328	0.3523	3.99	1.582	1.98
RYC 5/4	0.8113	0.5342	0.4381	5.61	2.047	2.41

Note: The first number next to the name of the strategy denotes the maturity of a fixed income investment whereas the second number stands for the holding period.

Table 32: Performance of the RYC Strategy (Conditional)

Strategy	Sharpe ratio	Excess return on VaR	Conditional Sharpe ratio	Omega	Sortino ratio	Upside potential ratio
RYC 2/1	0.4489	0.2387	0.2167	2.44	0.410	0.58
RYC 3/1	0.4787	0.2586	0.2284	3.10	1.122	1.48
RYC 3/2	0.4365	0.2306	0.2016	2.33	0.893	1.28
RYC 4/1	0.4742	0.2555	0.2173	3.04	1.145	1.52
RYC 4/2	0.4816	0.2606	0.2268	2.83	1.072	1.45
RYC 4/3	0.7149	0.4427	0.3106	8.42	2.262	2.53
RYC 5/1	0.4733	0.2549	0.2271	3.04	1.138	1.51
RYC 5/2	0.4919	0.2676	0.2350	3.00	1.122	1.50
RYC 5/3	0.4512	0.2401	0.2121	2.78	1.004	1.37
RYC 5/4	0.4466	0.2371	0.2003	2.81	1.012	1.37

Note: The first number next to the name of the strategy denotes the maturity of a fixed income investment, whereas the second number stands for the holding period.

Performance measures other than the Sharpe ratio confirm this result showing the same tendency to rise with the longer holding horizon. This finding contradicts the earlier results obtained for Germany in the study of Bieri and Chincarini (2005), who found the excess returns to increase with maturity of the riding instrument, but not with the holding period. Regarding the performance of the instruments of different maturities, the performance indicators tend to increase with higher maturities, keeping the holding period constant. An exception is the ride for one year: if a desired holding period constitutes one year, the best performance is achieved with a three-year instrument while the five-year instrument does the worst.

For conditional RYC strategies, the pattern concerning the holding periods is not so clear like for the unconditional strategies. For example, holding a three-year bond for one year yields a Sharpe ratio of 0.4787, whereas holding the same bond for two years results in a Sharpe ratio of only 0.4365. In contrast, for riding with a four-year bond, one improves the Sharpe ratio when riding during two years instead of one year and further improves it moving from the two-year to a three-year holding horizon. As for the choice of the best maturity for riding the yield curve, it is beneficial to choose a longer maturity instrument if the holding horizon comprises one or two years. If one intends to ride for three years, a better option is to choose a four-year rather than a five-year bond. Table 33 reflects the ranking of conditional as well as unconditional RYC strategies.

Table 33: Ranking of the RYC Strategies

Strategy	Sharpe ratio	Excess return on VaR	Conditio-nal Sharpe ratio	Omega	Sortino ratio	Upside potential ratio
RYC 5/4	1	1	1	2	2	2
RYC 4/3F	2	2	4	1	1	1
RYC 5/3	3	3	2	3	3	3
RYC 4/3	4	4	3	4	4	4
RYC 5/2	5	5	6	12	5	5
RYC 4/2	6	6	5	13	7	6
RYC 3/2	7	7	7	14	6	7
RYC 5/2 F	8	8	8	8	10	10
RYC 4/2 F	9	9	11	9	12	12
RYC 3/1 F	10	10	9	5	11	11
RYC 4/1 F	11	11	12	6	8	8
RYC 5/1 F	12	12	10	7	9	9
RYC 5/3 F	13	13	14	11	14	13
RYC 2/1 F	14	14	13	15	20	20
RYC 5/4 F	15	15	16	10	13	14
RYC 3/2 F	16	16	15	16	15	15
RYC 3/1	17	17	18	17	16	16
RYC 2/1	18	18	20	18	18	19
RYC 4/1	19	19	19	19	17	17
RYC 5/1	20	20	17	20	19	18

Note: The first number next to the name of the strategy denotes the maturity of a fixed income investment whereas the second number stands for the holding period; an “F” means that the strategy was based on a filter rule.

Whereas the ranking according to the Sharpe ratio completely equals the ranking assigned by the ERVaR measure, the ranking across other performance measures slightly differs. According to the Sharpe ratio, the ERVaR and the conditional Sharpe ratio, the best performing RYC strategy is the unconditional strategy of buying a five-year bond and selling it after four years. This is the strategy with the longest maturity and holding horizon. The worst performance was shown by the unconditional riding with a five-year bond for one year. On the top of the list are unconditional strategies, especially those with long maturities and holding horizons. The best among the conditional strategies is riding the yield curve for three years using a four-year instrument. According to the Omega, the Sortino ratio and the UPR, it is even on the first place; the rest of the measures assign it second to fourth place. This is the only long maturity and long holding period strategy that outperforms an analogical unconditional strategy. The filter is thus effective only on the shortest holding horizon: the conditional strategies that ride for one year with maturities of two, three, four and five years outperform the identical non-filtered strategies according to most of the considered performance indicators.

Thus, for a one-year holding horizon, one is better off when relying on a positive slope filter rule rather than riding all the time. However, for longer holding periods, it is generally not beneficial to use such a filter rule. Similar results regarding the usefulness of the positive slope filter rule were obtained in section 4.3.2 where RDYC strategies were considered. Moreover, this result is also in line with the findings obtained for the US market as well as the study of Bieri and Chincarini (2005), who could not confirm the ability of a positive slope filter rule to enhance returns on the German data. Table 34 contains the rank correlation coefficients for all RYC strategies.

Table 34: Rank Correlation of Performance Measures (RYC)

Strategy	Sharpe ratio	Excess return on VaR	Conditional Sharpe ratio	Omega	Sortino ratio
UPR	0.938	0.938	0.922	0.805	0.995
Sortino Ratio	0.937	0.937	0.911	0.808	
Omega	0.809	0.809	0.782	1.000	
Conditional SR	0.973	0.973	1.000		
Excess return on VaR	1.000				

The rank correlation across performance measures is at least 78 percent, the lowest value in case of the CSR and the Omega. The rank correlation of the Sharpe ratio with other performance measures is in the range of 80.9 percent (with Omega) to 100 percent (with excess return on VaR). The Sharpe ratio and the ERVaR exhibit the highest average rank correlation with the other measures of 94.3 percent while the Omega displaces the lowest average rank correlation of 83.6 percent. Thus, various performance measures display high rank correlation with each other, also in case of the RYC strategies. Consequently, the choice of a particular measure is not crucial for ranking the RDYC and RYC strategies in the considered data set. Figure 21 depicts the cumulative returns from RYC strategies with a holding horizon of one year and the corresponding BH strategy. Figure 22 plots the cumulative distribution function of returns from RYC strategies with holding horizons beyond one year and the corresponding BH strategy.

Figure 21: Cumulative Probability Distribution of RYC with One Year Holding Horizon

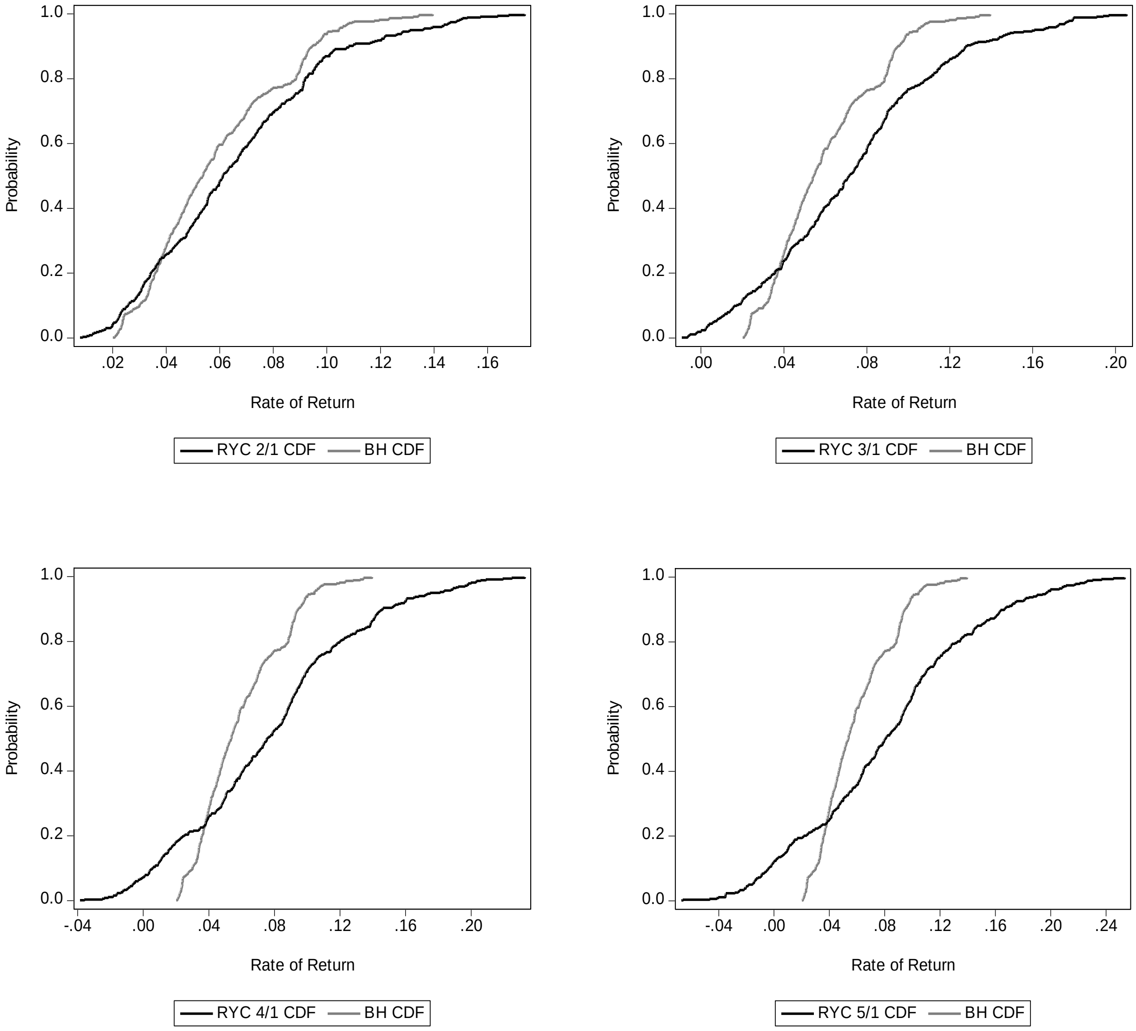
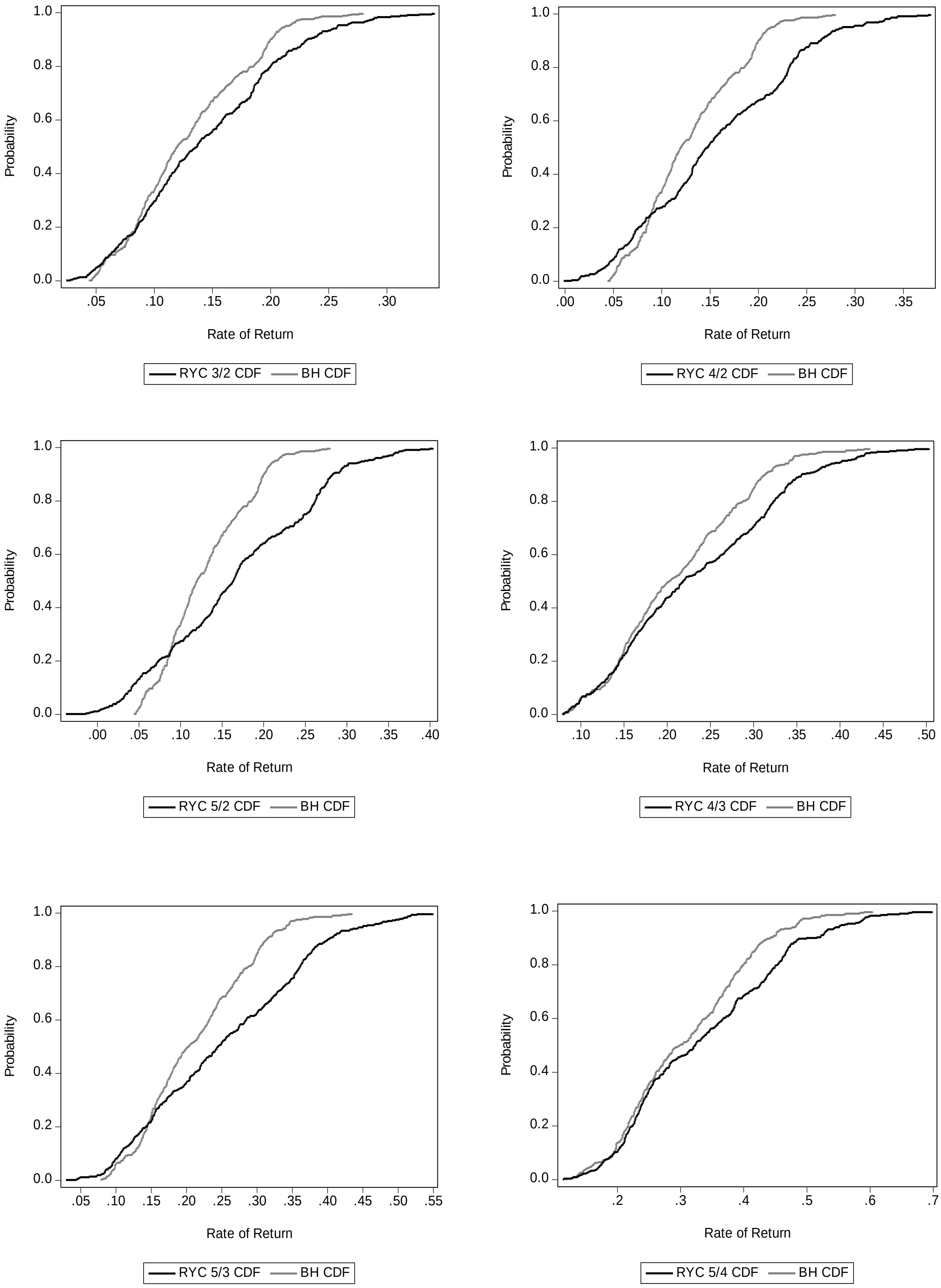


Figure 22: Cumulative Probability Distribution of RYC with Holding Horizon beyond One Year



As it is evident from figures 21 and 22, there is no first order stochastic dominance for all but one RYC strategies, as the respective cumulative distribution functions intersect. The intersection occurs at a low level of rate of return, and then the cdf of the RYC strategies lies below that of the benchmark strategy. The only exception is the strategy RYC 5/4 with the longest maturity instrument as well as holding period, whose cdf either lies below or is tangent to the cdf of BH strategy. Thus, RYC 5/4 dominates the corresponding BH strategy by the first order stochastic dominance. For the rest of the riding strategies, the area below the cdf is smaller than that of the BH strategy. Thus, RYC strategies dominate the corresponding BH strategy by the second order stochastic dominance criterion. These findings are in line with the study of Ang, Alles and Allen (1998), who could not confirm the first order stochastic dominance of the riding strategies, but found that the RYC strategies dominate the BH strategy in terms of the second order stochastic dominance.

4.6 Banking Regulation

Both the Rolling Down the Yield Curve and Riding the Yield Curve strategy described in the previous subsections possess an attractive risk-return profile, if compared to the DAX and the buy-and hold strategy, respectively. However, the risks attributable to these strategies can become significant. This section describes the main sources of risks and aims to give an overview over recent changes in the regulatory requirements, which could affect the possibility of implementation of the yield curve strategies. Section 4.6.1 provides an excursus into the regulatory framework and covers recent developments in regulatory requirements undertaken to promote financial stability. Readers familiar with these issues can proceed directly to section 4.6.2 that covers main sources of risk of the strategies developed in previous sections and addresses the main efforts to mitigate these types of risks.

4.6.1 Regulatory Framework: An Overview

Banking regulation and supervision in Germany was firstly established in 1931, with the publication of the Emergency Decree.[121] While there was no general regulation of the German

[121] “Decree relating to Stock Corporation Law, Banking Supervision and Tax Amnesty”.

banking sector prior to 1931,[122] the Emergency Decree was primarily targeted at overcoming the banking crisis of 1929–1930 and preventing future crises. Savings banks were not incorporated into this decree and were under the old regulatory rules.[123] Only in 1935, with the enforcement of the German Banking Act all banks were included. From 1949 and until the enactment of the Banking Act[124] and the foundation of the new supervisory authority in 1962, banking supervision had a decentralized character and was exercised by the respective state governments. With the extension of banking business activities, the necessity of regulatory adjustments became apparent, which led to several amendments to the Banking Act. Several revisions to the Banking Act were made in order to bring it in accordance with the new regulatory requirements of the Basel Community on Banking Supervision. The first version of these requirements, the Basel Capital Accord (known as Basel I), introduced the minimum capital requirements for financial institutions and was published in 1988. However, the limitations of this framework[125] led to several amendments[126] and, finally, to the new version of the Basel Capital Accord (known as Basel II)[127], which was published in 2004.

Pillar I of Basel II covers standard as well as advanced techniques for measuring credit, operational and market risks. Two further pillars include the supervisory review process (pillar II) and market discipline (pillar III). The former encourages banks to establish an efficient risk management system to assess their capital adequacy, whereas the latter contains public disclosure requirements, which enable market participants to assess the risk profile of a bank. The Basel II regulations constituted a foundation for the respective European Union (EU) directives, the Banking Directive and the Capital Adequacy Directive,[128] which were published in 2006. In order to create a legal basis for the implementation of Basel II in Germany, the new regulations were transferred into German law through amendments to the Banking Act, the Solvency Regulation, and the "Minimum Requirements for Risk Management" (MaRisk). Whereas pillars I and III of Basel II are reflected in the Solvency Regulation, pillar II is represented by the MaRisk. As a result, all banks in Germany are obliged to follow the

[122] Only savings banks and mortgage banks were regulated by the state governments and the federal government, respectively.
[123] See Hackethal/Schmidt (2005).
[124] The Banking Act is a legal basis for the supervision of the banking business and financial services in Germany.
[125] Basel I had been criticized because its method to calculate the regulatory capital for credit risk did not take into account the actual default risk of individual loans and, therefore, did not sufficiently reflect the actual credit risk.
[126] In 1996 the capital requirement to cover market risks was incorporated into the Basel Capital Accord.
[127] See BIS (2004).
[128] Directive 2006/48/EC and 2006/49/EC of June 14, 2006.

Basel II requirements since January 1, 2007.[129] Two regulatory authorities exercise banking supervision in Germany: the Federal Financial Supervisory Authority (FFSA) and the German Federal Bank.[130]

In the Basel II framework banks are required to hold capital, which is larger or equal to the eight percent of the risk-weighted assets for credit, market and operational risk. The regulation embodies two techniques to determine the credit risk regulatory capital, in particular the Standard Approach and the Internal Ratings-Based Approach (IRBA). In the framework of the Standard Approach, banks are allowed to determine risk weights on the basis of external credit ratings issued by a rating agency approved by the supervisor. At the same time, it is still possible to apply uniform fixed risk weights, which is especially relevant for unrated loans. The IRBA provides an opportunity to apply internal rating procedures and completely or partially estimate the input parameters for measuring credit risk.[131]

The Basel II regulations recognize operational risks and permit banks to choose among three different techniques to determine the capital charge for operational risk: the Basic Indicator Approach (BIA), the Standardized Approach (SA), and the Advanced Measurement Approach (AMA). Under the BIA, the capital charge is determined as the weighted average gross income of a bank over the past three years, which serves as an operational risk indicator. The SA requires this indicator to be broken down into eight business lines determined in the Solvency Regulation and multiplied by the weights of the business lines, which range from 12 to 18 percent. In contrast, in the AMA, which requires a prior approval of the FFSA, banks can calculate the capital requirements for operational risk using an internal model.[132]

Market risks include interest rate, equity position, exchange rate fluctuation and commodities risk. Basel II specifies two methods of measurement of such risks: the Standardized Measurement Approach and the Internal Models Approach. The latter is subject to supervisory approval. The market risk measure is VaR,[133] which has to be computed over the period of ten days using the 99 percent "confidence" interval. The observation period constitutes one year.

[129] See Reichling/Afanasenko (2010), pp. 13-14.
[130] The German expression of the German Federal Bank and the FFSA is Deutsche Bundesbank and Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin), respectively.
[131] See Reichling/Afanasenko (2010), pp. 14-15.
[132] See Reichling/Afanasenko (2010), p. 15.
[133] The choice of a particular method to compute VaR is not specified by Basel II and can be made by financial institutions. Variance-covariance approach, historical or Monte Carlo simulation could be applied.

In order to determine a capital requirement for market risk, the VaR figure is multiplied by a factor which is set up by the regulatory authorities. It has a minimum value of three and can be raised, dependent on the past performance of the corresponding model.[134]

As a response to the financial crisis in 2008, the Basel Committee on Banking Supervision was forced to revise its regulatory framework. In December 2010 a new document, Basel III[135] was issued, which was revised in June 2011. The main feature of the new framework is the new qualitative and quantitative requirements towards the capital base. Although the amount of total capital[136] stays at the level of eight percent, the structure of total capital will undergo significant changes. First of all, total capital is now comprised of only two categories: Tier I and Tier II, i.e. the earlier division of total capital into three categories will be abolished. Tier I capital is further divided into two sub-categories: common equity Tier I and additional Tier I capital. The amount of a high-quality common-equity Tier I capital will be raised from two percent to 4.5 percent whereas additional Tier I capital will be reduced from two percent to 1.5 percent. Thus, the total amount of Tier I capital will increase from four percent to six percent. As for the Tier II capital, it will be reduced from four percent to only two percent. The above transitions will be implemented gradually and completed by January 2019.[137]

Another feature of the Basel III regulations is the incorporation of two buffers: the capital conservation buffer and the countercyclical buffer, which are to be held above the minimum capital requirements. The former is to be used in crisis times and serves as an additional capital that can be drawn down without consequences for the minimum capital requirements or restrictions of business activities of the financial institutions. However, if this buffer is completely depleted, the bank will encounter restrictions regarding share buy-backs, dividends and staff bonus payments.[138] The amount of this capital, which must be held as common equity Tier I capital, will constitute 0.625 percent in 2016 and will gradually increase to reach its maximum of 2.5 percent in 2019.[139]

[134] See Bank for International Settlements (2004), pp. 195-197.
[135] See Bank for International Settlements (2010) and (2011).
[136] Total capital refers to the regulatory capital and not the balance sheet total.
[137] See Bank for International Settlements (2011), pp. 12-20 and pp. 27-29.
[138] This regulation was introduced to avoid practices of some banks that, despite losses, continued distributions of earnings to demonstrate their positive performance.
[139] See Bank for International Settlements (2011), pp. 54-57.

The second type of buffer – the countercyclical buffer – aims to reduce the pro-cyclicality of previous regulations. The essence of this problem is that excess credit growth in boom times can result in substantial losses for the banking sector if economic conditions change rapidly. In a recession, banks are forced to reduce the amount of granted loans and, thus, contribute to further deterioration of economic conditions. The new buffer is intended to be used in times of economic downturn and contribute to a faster recovery of the economy. The supervisory authorities may demand the amount of the countercyclical buffer to be raised if there is a need to do so in order to address the current economic conditions. Similarly to the capital conservation buffer, the countercyclical buffer will be implemented in the period 2016 – 2018 and will gradually increase from 0.625 percent to 2.5 percent.[140]

An additional characteristic feature of the new regulations is the leverage ratio. Despite strong capital ratios, many banks built up excessive leverage in the times of the financial crisis. As a result, banks were forced to reduce the amount of assets in a relatively short time, which put some additional pressure on asset prices. The aim of the new non-risk based leverage ratio is to supplement the existing risk-based capital requirements and restrict the level of leverage. During a testing period from 2013 to 2017, this ratio will constitute three percent of total assets. After that, the Basel Committee will decide, whether this ratio will become a part of the regulatory capital requirements.[141] In addition to the above mentioned developments in the regulatory framework, the Financial Stability Board (FSB) issued a new framework regarding the systemically important financial institutions (SIFIs).[142] The main focus of this document is the moral hazard problem that occurs in the institutions, whose bankruptcy would endanger the stability of financial systems. Such institutions will have to develop individual financial restructuring procedures and can be subject to additional capital requirements. The next section addresses the specific risks of the yield curve strategies and the corresponding regulations.

4.6.2 Specific Risks of the Yield Curve Trading Strategies

Rolling down the yield curve is an especially attractive strategy in a low interest rates environment. If banks are able to borrow short-term funds at a very low interest rate and grant long-term credits, potential profits are quite large. However, the strategy should be imple-

[140] See Bank for International Settlements (2011), pp. 57-60.
[141] See Bank for International Settlements (2011), pp. 61-63.
[142] See FSB (2010).

mented with care, as risks arising from the strategy could become significant. At first, the strategy is connected with interest rate risk. The short-term interest rate that will prevail in the future is not known and could increase so significantly, that short-term borrowing would become more expensive than the interest rate obtained on the long-term funds. Because of the high refinancing costs, the excess return on the strategy could then become negative. Various risk measures presented in table 35 indicate that risks can become significant, especially when using the strategy over several years. Looking at the VaR measure computed for pursuing the strategy for the period 1978 – 2007, the maximum negative excess return to occur with 99 percent probability is as high as 17.74 percent for a five-year strategy. In contrast, implementing the strategy for two years yields a VaR of 3.55 percent. If CVaR is employed as an interest rate risk measure, the average loss exceeding the VaR varies from 5.04 percent for a two-year strategy and 19.13 percent for a five-year strategy. For RDYC strategies with filter rule, the two-year strategy has a VaR of only 2.74 percent, which is lower than that of a corresponding non-filtered strategy. However, the five-year strategy implemented using a filter happens to be the riskiest of all, with VaR comprising 18.46 percent. CVaR lies in the range 3.19 percent to 19.39 percent.

Table 35: VaR and CVaR of the RDYC Strategy (in percent)

	RDYC 2	RDYC 3	RDYC 4	RDYC 5	RDYC 2F	RDYC 3F	RDYC 4F	RDYC 5F
VaR	3.55	7.86	12.85	17.74	2.74	7.41	12.54	18.46
CVaR	5.04	8.42	14.14	19.13	3.19	8.18	13.84	19.39

Similarly to the RDYC strategy, the RYC strategy is subject to interest rate risk. The strategy is constructed so that a fixed income instrument is bought with a maturity *m*, which is longer than the desired holding period *i*. Thus, it involves selling the instrument prior to its maturity, after $m - i$ periods. The price of the bond and, therefore, the profit from the strategy is highly dependent on the interest rate on the bond of maturity $m - i$ prevailing in the time of a bond sale. If this interest rate rises sharply, banks can experience a negative excess return on the strategy, compared to an initial purchase of a bond with the maturity *i*. Table 36 contains the VaR as well as CVaR of various RYC strategies implemented over the period of 30 years. The VaR and CVaR constitute 3.3 percent and 4.48 percent for riding a two-year bond for one year. The RYC strategy that was especially prone to interest rate risk was riding a five-year bond for two years, whose VaR and CVaR comprised 11.46 percent and 13.19 percent, respectively.

Table 36: VaR and CVaR of the RYC Strategy (in percent)

	RYC 2/1	RYC 3/1	RYC 3/2	RYC 4/1	RYC 4/2	RYC 4/3	RYC 5/1	RYC 5/2	RYC 5/3	RYC 5/4
VaR	3.30	6.18	4.64	8.59	8.43	5.12	10.62	11.46	9.41	5.25
CVaR	4.48	8.40	5.28	11.37	9.35	6.55	13.84	13.19	11.57	6.41

The second type of risk that has to be accounted for when implementing the RDYC strategy is the liquidity risk. Liquidity risk refers to the risk that the liabilities cannot be met. Considering the RDYC strategy, this type of risk arises because the maturity of assets does not match the maturity of the liabilities. A bank that pursues the RDYC strategy faces liquidity risks that occur when it has to refinance itself when the initial short-term funding has to be paid pack. The fact that this risk can become significant was especially obvious in fall 2008. In this time, many financial institutions encountered liquidity shortage. These difficulties were so severe and occurred in such a large number of financial institutions that the interbank market collapsed. As a result, central banks were forced to replace the interbank market.

The regulators became aware of the fact that liquidity risks can play a significant role for the stability of financial systems and introduced new regulatory requirements. They aim at stressing the liquidity base of financial institutions and reducing their dependence on the short-term financing from the interbank market. At first, the Basel Committee introduced new standards regarding qualitative requirements for liquidity risk management.[143] Later, in December 2010, the Basel Committee issued new guidance that introduced not only qualitative, but also quantitative global liquidity standards.[144] The new framework introduced two measures of liquidity risk exposure, the Liquidity Coverage Ratio (LCR) and the Net Stable Funding Ratio (NSFR). Under various stress scenarios, the former should enable banks to survive within 30 days relying only on their own sources of funding. The LCR, which is set to be implemented by 1 January, 2015, is given by:[145]

[143] See Bank for International Settlements (2008).
[144] See Bank for International Settlements (2010).
[145] See Bank for International Settlements (2010), p. 3.

$$\text{LCR} = \frac{\text{Stock of high-quality liquid assets}}{\text{Total net cash outflows over the next 30 calendar days}} \geq 100\% \tag{85}$$

Only a limited number of assets is eligible to be counted as high-quality liquid assets. Such assets include cash, reserves by central banks as well as government securities and constitute so called level 1 assets. Corporate bonds of high credit quality belong to the level II assets, which are allowed to comprise no more than 40 percent of total stock of high-quality liquid assets.[146] The NSFR represents a long-term liquidity standard, which will become obligatory on 1 January, 2018. Over a one-year horizon, banking institutions must maintain the following ratio:[147]

$$\text{NSFR} = \frac{\text{Available amount of stable funding}}{\text{Required amount of stable funding}} \geq 100\% \tag{86}$$

The available amount of stable funding is defined as the sum of all sources of funding, weighted by the Available Stable Funding (ASF) factor. The weights vary between zero percent and 100 percent; for example, Tier I and II capital receives the weight of 100 percent whereas stable demand deposits of retail and small business customers are accounted for using a 90 percent weight. It is worth noting that the possibility of refinancing through central banks is not accounted for in the numerator of equation (86). Otherwise, banks could rely too strongly on this source of funding. The required amount of stable funding is determined as a weighted sum of all assets. The weights ranging from zero to 100 percent reflect the liquidity of various asset types, i.e. a possibility to liquidate the assets within one year. Assets with a low necessity to be covered by stable funding, such as cash and securities with a maturity of less than one year have a weight of zero percent whereas loans to retail customers that mature in less than one year receive a weight of 85 percent.[148]

The NSFR is set to limit the incentives of financial institutions to establish a strong mismatch between the maturity of the assets and liabilities and is, therefore, especially important in the context of the RDYC strategy. The implementation of the RDYC strategy where long-term assets are funded through liabilities that mature in less than one year will be restricted by the

[146] See Bank for International Settlements (2010), pp. 8-10.
[147] See Bank for International Settlements (2010), pp. 25-26.
[148] See Bank for International Settlements (2010), pp. 26-29.

NSFR. The intention of this part of liquidity regulation is to reduce the over-reliance on short-term funding at the interbank market. However, the fact that the ratio is applied only during a one-year horizon indicates that the implementation of the RDYC strategy, which is an especially important source of income for some banking groups in Germany, will be further possible.

5. Summary

The focus of this dissertation thesis was on analyzing the predictive power of forward rates as well as examining the performance of various yield curve trading strategies. The pure expectations hypothesis of the term structure of interest rates provided the basis for the first part of the analysis. Among the classical term structure theories that include the pure expectations theory, the liquidity preference, the preferred habitat theory and the market segmentation theory, the pure expectations theory received the most attention in the academic literature. The reason is that this version states that long-term interest rates represent the geometric average of the current and expected future short-term interest rates and is, therefore, appealing due to its simplicity. An additional benefit offered by the theory is that it can explain every possible shape of the term structure. An upward-sloping term structure results from the situation where investors anticipate future interest rates to rise. Similarly, a downward-sloping as well as a flat term structure can be explained.

The theory also addresses two further empirical facts usually associated with the term structure: the first one is that interest rates of different maturities tend to move together; the second fact is that short-term interest rates are more volatile than the long-term interest rates. Both observations are in line with the pure expectations hypothesis, as long-term rates are an average of the short-term rates and thus, they should not develop independently from each other. The main drawback of the pure expectations hypothesis lies in its inability to explain the fact that, on average, the yield curve is upward-sloping.

The implications of the pure expectations hypothesis were empirically tested in a variety of ways. However, the general result is that the theory is not supported by the empirical data. In Germany, there have been several studies on that topic; yet leading to contradictory results. The research presented in this dissertation thesis aims to investigate the predictive power of forward rates towards the one-year spot interest rates using recent data and new estimation techniques. At the first glance, forward rates seem to systematically overestimate the future spot rates; this overestimation rises with the lag of the respective forward rate. However, a formal analysis was needed, in order to achieve reliable results.

In this dissertation thesis, six models were constructed. Whereas the first model explains the one-year spot rate by the corresponding forward interest rate that was observed one year ago, the last model contains six lagged forward rates. The spot rates employed in this study were computed using the swap rates for the period 1995 – 2007. For the choice of the estimation method, it was indispensable to test the time series properties of forward as well as of spot rates. It was especially important to examine, whether considered series possess the necessary statistical properties that make standard regression analysis possible. This property is referred to as the stationarity of the time series. The stationarity tests employed in this study include the augmented Dickey-Fuller, Philips-Perron, Dickey-Fuller generalized least squares and Kwiatkowski-Philips-Schmidt-Shean test. Their results were quite uniform, indicating that spot rates as well as forward rates are non-stationary time series that become stationary after taking first differences. In other words, spot and forward rates are non-stationary, integrated of order one series.

Because of the special time series properties, standard regression techniques are not applicable, as the results could possibly indicate a spurious relation. However, a meaningful relationship between two time series exists if they are cointegrated. The concept of cointegration is useful, as it refers to the series which are individually non-stationary, but a linear combination can be found between them that is stationary. The Johansen cointegration test performed for the one-year spot rate and lagged forward rates confirmed that the series are cointegrated and the number of cointegrating vectors is equal to one for all six models. This can be viewed as evidence that the long-term relationship among spot rate and forward rates exists. This result is in line with several previous studies which documented the presence of cointegration in the term structure of interest rates.

Despite the fact that the one-year spot interest rate and lagged forward rates form a cointegration relationship, reliable evidence that forward rates can be used as predictors of the future spot rates was not found. In the majority of the considered models only the forward rate observed one year ago had significant impact on the future spot rate. The six-year model seems to explain the data the best in this respect containing the largest number of significant forward rates. Commonly for all considered models, only the one-year, the five-year and the six-year forward rates are statistically significant and, therefore, affect the future spot rates. This could be interpreted as an evidence of a mean-reverting behavior; however, this hypothesis would contradict the results of earlier stationarity tests about non-stationarity, and, therefore, an absence of constant mean and variance in the considered data.

In addition to the cointegration analysis, an error correction representation was constructed to capture the short-run dynamics. The adjustment coefficient plays an important role, as its significance and sign provides some insights about how well the Error Correction Model fits the data. The estimated error correction representation has a significant adjustment coefficient in all considered models, which gives an evidence of the validity of the Error Correction Model in the data. However, the sign of this coefficient is negative only for the five-year and six-year models. As a negative sign is essential for the Error Correction Model to work (i.e. to correct the error which has occurred in the previous period), the Error Correction Model is valid only in case five or six lagged forward rates are included into the model. In addition, the Error Correction Model has a poor fit, with the coefficient of determination ranging from ten percent for the two-year model to the maximum of 23 percent for the six-year model.

As a concluding part of this analysis, it was of interest to check the out-of-sample performance of the models. For this reason, the data for the last available year were not used to estimate the required coefficients. The estimates of the cointegration equation obtained with the sample May 1995 to October 2006 were employed to construct a forecast for the next 12 months. The resulting forecasts of each of the six models were compared to a simple model, which just uses the spot rate of the previous period to predict the future spot rate. The forecasting performance of these models was then evaluated with the help of commonly applied forecast accuracy measures, such as root mean squared error and Theil's inequality coefficient. These measures indicate that the one-year to five-year models are not really worth the effort of estimating them, as using the last period's value of one-year spot rates yields a lower forecast error. Only the six-year model performs better than the naive model. This reinforces the conclusion derived from the cointegration and error correction analysis. The general result of the analysis is that forward rates contain very poor predictive ability and generally cannot serve as predictors of future spot rates.

Thus, the results of the first part of the analysis provide additional evidence against the validity of the Expectations Hypothesis in the German term structure. Although several studies have found some supportive evidence supporting the validity of the Expectations Hypothesis in German data, the findings in this thesis are in line with those studies that were not able to find the information content in forward rates. An important implication can be derived from this result: forward rates are not reliable predictors of the future spot rates and, thus, the fact that forward rates exceed the spot rates most of the time does not necessarily indicate rising spot rates in the future. In other words, if the pure Expectations Hypothesis does not hold, holding period returns of assets with different maturities may not be the same.

The second part of this thesis deals with two yield curve strategies: Rolling Down the Yield Curve and Riding the Yield Curve. A common feature of these strategies is that they are both based on an upward-sloping yield curve which remains stable over time. The Rolling Down the Yield Curve strategy is a strategy commonly used by financial institutions and can be described as funding a long-term asset using a short-term liability. Such a strategy can be implemented, for instance, as a typical banking business of granting long-term loans and taking short-term deposits. In addition, the strategy can be constructed through buying and selling fixed income securities or through activities on the swap market. This part of the analysis aims to determine the risk-return profile of the Rolling Down the Yield Curve strategy and compare it with that of the DAX. The latter represents the benchmark for investments on the capital market.

Four types of the Rolling Down the Yield Curve strategy were examined: as a short-term funding a one-year horizon is selected, whereas the long-term investment varies from two to five years. In addition, every of the above strategies was tested with and without a statistical filter rule. Strategies where the filter rule was not applied were just implemented all the time, independently of the shape of the term structure. Those strategies conditioned on a filter were only pursued in case the term structure of interest rates was upward-sloping. The period under consideration covers 1972 – 2007.

Already the visual representation of the excess return-standard deviation profile makes obvious that the two-year strategy dominates the DAX, as it has both higher excess return and lower standard deviation. For the rest of the strategies the risk-return profile does not deliver such clear results; thus, various performance indicators have been applied and the strategies were ranked dependent on their performance. Among the employed performance measures are traditional measures such as the Sharpe ratio as well as modern measures based on value at risk or lower partial moment. The issue, whether a particular performance measure is appropriate to be used for ranking, received significant attention in the recent literature. In this dissertation thesis, all most commonly applied measures were included into the analysis. Although the ranking slightly differs among performance measures, a common feature is that all Rolling Down the Yield Curve strategies clearly outperform the DAX. This is an important result, as it actually indicates the success of client business activities over the fee-based activities of banks on the capital market.

Regarding the performance of particular strategies, the Rolling Down the Yield Curve strategy implemented through an investment in a five-year instrument showed the best performance. In general, the strategy becomes more attractive with the length of the implementation period. Although both excess returns and risk rise, strategies implemented for a longer period possess a more attractive risk-return tradeoff. The applied filter rule, however, was useless, as it did not yield superior performance of the filtered strategies. In fact, all Rolling Down the Yield Curve strategies where such a filter was applied performed worse than the strategies that were implemented all the time. Thus, only relying on an upward-sloping yield curve when the strategy is initiated does not guarantee a success of the strategy and even worsens the performance.

An interesting insight also provides a comparison of rankings assigned by different performance measures. The results in this thesis indicate that, although the ranking of strategies slightly varies among performance measures, the rank correlation coefficients are quite high for each pair of measures. Thus, the choice of a particular performance indicator did not matter in this study.

It is worth noting that Rolling Down the Yield Curve is not an arbitrage strategy. In fact, the strategy is subject to interest rate risk. If, contrary to the anticipations, the term structure shifts upwards or flattens out, short-term interest rates could rise so significantly that banks suffer losses from the strategy. Although it would be possible to apply some hedging techniques to eliminate risks, this would eliminate profits from the strategy. From a theoretical point of view, it is only possible to gain profits above the risk-free rate of return if an investment contains risks. Otherwise, the rate of return would equal the risk-free rate of return. Thus, the Rolling Down the Yield Curve strategy could be referred to as “risk arbitrage”, as it provides significantly more attractive risk-return tradeoff than the capital market investment.

Despite the risks associated with the strategy, it proved to be successful over the considered period of more than 30 years. In comparison with the investment on the stock market, the Rolling Down the Yield Curve strategy represents a more attractive business activity. This result is of a special importance for banking groups whose main sources of income constitutes the interest rate income.

The second strategy examined in this thesis was the Riding the Yield Curve strategy. Similarly to the Riding the Yield Curve strategy, it is based on an upward-sloping term structure that remains constant over time. The Riding the Yield Curve strategy involves buying an instrument with a maturity longer than the anticipated holding horizon and selling it prior to maturity. It is regarded as a contradiction to the postulates of the pure expectations theory. As the results of the first part of the analysis did not support the hypothesis, analyzing the performance of the Riding the Yield Curve strategy could provide further confirmation of the previously obtained results. The strategy was implemented in using maturities from two to five years and holding periods from one to four years. In addition, every strategy was implemented with a filter conditioned on a positive slope of the yield curve. Every Riding the Yield Curve strategy under consideration was compared with buying and holding an instrument till maturity.

It follows from the analysis that all Riding the Yield Curve strategies outperform the simple buy-and-hold strategy, as they yield positive average excess returns. The results of the performance evaluation indicate that riding the longest maturity instrument during the longest possible period proved to be the most beneficial. In contrast, investing long-term and selling after one year was the strategy with the worst performance. As in case of the Rolling Down the Yield Curve strategy, the application of the filter rule generally did not lead to a superior performance of the Riding the Yield Curve strategies. The latter result is in line with some previous studies devoted to the Riding the Yield Curve strategy. The superior performance of the Riding the Yield Curve strategy once more confirms previous results on the validity of the pure expectations hypothesis. Not only are forward rates poor predictors of the future spot rates, but there is also a possibility to gain excess returns from riding the yield curve. Thus, predictions relying on forward rates should be treated with caution.

Finally, the risks connected with the yield curve strategies should not be underestimated. Except of the interest rate risk, the Rolling Down the Yield Curve strategy is subject to liquidity risk. The key problem is that it involves refinancing its long-term assets through short-term liabilities. The situation in which a bank experiences severe difficulties in attracting short-term funds could even endanger the existence of this financial institution. This was especially obvious during the liquidity crisis in 2008. New regulatory requirements proposed in Basel III aim to reduce these risks by introducing new quantitative standards that will become obligatory on 1 January, 2018. Whereas the Liquidity Coverage Ratio is introduced to ensure the survival of a financial institution during a 30-day period on its own, the Net Stable Funding Ratio is targeted at reducing the maturity mismatch between assets and liabilities over a one-year horizon. The latter, although reducing the possibilities of reliance on the short-term funding, still leaves some space to banks for pursuing the strategy.

References

Acerbi, C., Nordio, C., and Sirtori, C. (2001): Expected Shortfall as a Tool for Financial Risk Measurement, Working Paper.

Afanasenko, D., Gischer, H., and Reichling, P. (2011): The Predictive Power of Forward Rates, *Investment Management and Financial Innovations* 8, 125–139.

Agarwal, V., Naik, N.Y. (2004): Risk and Portfolio Decisions Involving Hedge Funds, *Review of Financial Studies* 17, 63–98.

Akaike, H. (1974): A New Look at Statistical Model Identification, *IEEE Transactions on Automatic Control* 19, 716–723.

Alexander, G.J., Sharpe, W.E. and J.V.Bailey (2001): Fundamentals of Investments, 3rd ed., Prentice Hall, Upper Saddle River, New Jersey.

Andrews, D. (1993): Tests for Parameter Instability and Structural Change with Unknown Change Point, *Econometrica* 61, 821–856.

Ang, S., Alles, L. and Allen, D. (1998): Riding the Yield Curve: An Analysis of International Evidence, *Journal of Fixed Income* 8, 57–74.

Ang. A. and Piazessi, M. (2003): A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, *Journal of Monatery Economics* 50, 745–787.

Artzner, P., Delbaen, F., Eber, J-M., Heath, D. (1997): Thinking Coherently, *Risk* 10, 68–71.

Artzner, P., Delbaen, F., Eber, J-M., Heath, D. (1999): Coherent Measures of Risk, *Mathematical Finance* 9, 203–218.

Asteriou, D. and Hall, S.G. (2006): *Applied Econometrics: A Modern Approach using E-Views and Microfit*, Palgrave Macmillan, New York.

Bank for International Settlements (2008): Principles for Sound Liquidity Risk Management and Supervision, http://www.bis.org/pub/bcbs144.htm, accessed on 15 December 2008.

Bank for International Settlements (2010): Basel III: International Framework for Liquidity Risk Measurement, Standards and Monitoring.

Bank for International Settlements (2011): Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems.

Bawa, V. (1975): Optimal Rules for ordering Uncertain Prospects, *Journal of Financial Economics* 2, 95–121.

Bawa, V. (1978): Safety first, stochastic dominance, an optimal portfolio choice, *Journal of Financial and Quantitative Analysis* 13, 255–271.

Beechey, B. Hjalmarsson, E. and Oesterholm, P. (2009): Testing the Expectations Hypothesis when Interest Rates are Near Integrated, *Journal of Banking and Finance* 33, 934–943.

Benning, J.F. (2007): *Trading Strategies for Capital Markets*, McGrow-Hill, New York.

Bieri, D.S. and Chincarini, L. (2005): Riding the Yield Curve: A Variety of Strategies, *Journal of Fixed Income* 15, 6–35.

Bodie, Z., Kane, A., Marcus, A.J. (1999): *Investments*, 4th ed., McGrow-Hill, Boston.

Brooks, C. (2008): *Introductory Econometrics for Finance*, 2nd ed., Cambridge University Press.

Boero, G. and Torricelli, C. (2002): The Information in the Term Structure of German Interest Rates, *European Journal of Finance* 8, 21–45.

Burke, S.P. and Hunter, J. (2005): *Modelling Non-Stationary Time Series*, Palgrave Macmillan, New York.

Campbell, J.Y. and Shiller, R.J. (1991): Yield Spreads and Interest Rates Movements: A Bird's Eye View, *Review of Economic Studies* 58, 495–514.

Chandy, P.R. and Hsueh, P. (1995): Riding the Yield Curve: An Empirical Examination of Recent US Data, *Singapore Management Review* 17, 31–39.

Chatfield, C. (2001): *Time-Series Forecasting*, Chapman & Hall, Boca Raton.

Chen, R.R. and Scott, L. (1993): A Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates, *Journal of Fixed Income* 3, 14–31.

Chow, G.C. (1960): Test of Equality between Sets of Coefficients in Two Linear Regressions, *Econometrica* 28, 591–605.

Cooray, A. (2003): A Test of the Expectations Hypothesis of the Term Structure of Interest Rates for Sri Lanka, *Applied Economics* 35, 1819–1827.

Crouhy, M., Galai, D., Mark, R. (2009): *The Essential of Risk Management*, McGrow-Hill, New York.

Cuthbertson, K., and Nitzsche, D. (2004): *Quantitative Financial Economics*, 2nd ed., Wiley, West Sussex.

Dai, Q. and Singleton, K. J. (2000): Specification Analysis of Affine Term Structure Models, *Journal of Finance* 55, 1943–1978.

Danielsson, J. (2002): The Emperor Has No Clothes: Limits to Risk Modeling, *Journal of Banking and Finance* 26, 1273–1296.

Darst, D.M. (1975): *The Complete Bond Book*, McGraw Hill, New York.

Davidson, J.E.H., Hendry, D.F., Shrba, F. and Yeo, S. (1978): Economic Modeling of the Aggregate Time Series Relationship Between Consumers' Expenditure and Income in the United Kingdom, *Economic Journal* 88, 661–692.

DeJong, D.N., Nankervis, J.C., Savin, NE. and Whiteman, C.H. (1992): Integration versus Trend Stationary in Time Series, *Econometrica* 60, 423–433.

De Leonardis, N.J. (1966): Opportunities for Increasing Earnings on Short-Term Investments, *Financial Executive*, July, 48–53.

Den Butter, F.A.C. and Jansen, P.V. (2004): An Empirical Analysis of the German Long-Term Interest Rate, *Applied Financial Economics* 14, 731–741.

Dickey, D.A. and Fuller, W.A. (1979): Distributions of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association* 74, 427–431.

Dominguez, E. and Novales, A. (2002): Can Forward Rates be Used to Improve Interest Rate Forecasts? *Applied Financial Economics* 12, 493–504.

Dowd, K. (2000): Adjusting for Risk: An Improved Sharpe Ratio, *International Review of Economics and Finance* 9, 209–222.

Dyl, E. and Joehnk, D. (1981): Riding the Yield Curve: Does it Work? *Journal of Portfolio Management*, Spring, 13–17.

DZ Bank (2005): Zinsmärkte – Entwicklung einer Datenhistorie für die EUR Swap-Kurve, Research-Publikation, 20.12.2005.

Eling, M. and Schuhmacher, F. (2007): Does the Choice of Performance Measure Influence the Evaluation of Hedge Funds? *Journal of Banking and Finance* 31, 2632–2647.

Eling, M. (2008): Does the Measure Matter in the Mutual Fund Industry? *Financial Analyst Journal* 64, 54–66.

Elliot, G., Rothenberg, T.J and Stock, J.H. (1996): Efficient Tests for an Autoregressive Unit Root, *Econometrica* 64, 813–836.

Embrechts, P., McNeil, S. and Straumann, D. (2002): Correlation and Dependence in Risk Management: Properties and Pitfalls, in Dempster M.A.H. (ed.), *Risk Management: Value at Risk and Beyond*, Cambridge University Press, Cambridge.

Engle, R.F. and Granger, C.W. (1987): Cointegration and Error-Correction: Representation, Estimation and Testing, *Econometrica* 55, 251–276.

Engsted, T. and Tanggaard, C. (1994): Cointegration and the US Term Structure, *Journal of Banking and Finance* 18, 167–181.

Evans, C.L. and Marshall, D.A. (2002): Economic Determinants of the Nominal Treasury Yield Curve, Federal Reserve Bank of Chicago, Working Paper.

Fabozzi, F.J. (2004): *Bond Markets, Analysis, and Strategies,* 5th ed., Pearson Education, Upper Saddle River, New Jersey.

Fabozzi, F. and Mann, S. (2010): Introduction to Fixed Income Analytics, 2nd ed., Wiley, New Jersey.

Fama, E. (1976): Forward Rates as Predictors of Future Spot Rates, *Journal of Financial Economics* 13, 509–528.

Fama, E. (1984): The Information in the Term Structure, *Journal of Financial Economics* 13, 509–528.

Fama, E. and Bliss, R. (1987): The Information in Long Maturity Forward Rates, *American Economic Review* 77, 680–692.

Favre, L. and Galeano, J.A. (2002): Mean-Modified Value at Risk Optimization with Hedge Funds, *Journal of Alternative Investments* 5, 21–25.

Feibel, B.J. (2003): *Investment Performance Measurement*, Wiley, New Jersey.

Fishburn, P.C. (1977): Mean-variance Analysis with Risk Associated with Below-Target Returns, *The American Economic Review* 67, 116–126.

Fisher, I. (1896): Appreciation and Interest, *The American Economic Review* 11, 331–442.

FSB (2010): Reducing the Moral Hazard Posed by Systemically Important Financial Institutions, http://www.financialstabilityboard.org/publications/r_101111a.pdf, accessed on 17 October 2010.

Flores, R.G. and Szafarz, A. (1996): An Enlarged Definition of Cointegration, *Economic Letters* 50, 193–195.

Frankfurter Wertpapierbörse (2011): Preisverzeichnis, www.deutsche-börse.com, accessed on 23 May 2011.

Freund, W.C. (1970): *Investment Fundamentals*, The American Bankers Association.

Frey, R. and McNeil A.J. (2002): Value at Risk and Expected Shortfall in Portfolios of Dependent Credit Risks: Conceptual and Practical Insights, *Journal of Banking and Finance* 26, 1317–1334.

Gerlach, S. and Smets, F. (1997): The Term Structure of Euro-Rates: Some Evidence in Support of the Expectations Hypothesis, *Journal of International Money and Finance* 16, 305–323.

German Federal Bank (1997): Estimating the Term Structure of Interest Rates, Monthly Report, October 1997.

Ghazali, N.A. and S.-W. Low (2002): The Expectations Hypothesis in Emerging Financial Markets: the Case of Malaysia, *Applied Economics* 74, 116–126.

Gischer, H. (1997): The Information Content of Forward Rates – Empirical Evidence from Germany; in: Hipp, C. et al. (ed.), *Geld, Finanzwirtschaft, Banken und Versicherungen*, Karlsruhe, 243–256.

Gischer, H. (1998): Der Erklärungsgehalt impliziter Terminzinsen – Ergänzungen zu Adam, Hering und Johanwille, *Zeitschrift für Betriebswirtschaft* 34, 1147–1156.

Granger, C.W. and Newbold, P. (1974): Spurious Regressions in Econometrics, *Journal of Econometrics* 14, 111–120.

Gregoriu, G.N., and Gueyie, J.P. (2003): Risk Adjusted Performance of Funds of Hegde Funds Using a Modified Sharpe Ratio, *Journal of Wealth Management* 6, 77–83.

Grieves, R., and A.J. Marcus, (1992): Riding the Yield Curve: A Reprise, *Journal of Portfolio Management* 18, 67–76.

Grieves, R., Mann, S., Marcus, A.J. and Ramanlal, P. (1999): Riding the Bill Curve, *Journal of Portfolio Management* 25, 74–82.

Gonzalez, J., Spencer, R. and Walz, D. (1999): The Information in the Mexican Term Structure of Interest Rates: Capital Market Implications, *Journal of International Financial Markets, Institutions and Money* 9, 149–161.

Guest, R. and McLean, A. (1998): New Evidence on the Expectations Theory of Australian Commonwealth Government Treasury Yields, *Applied Financial Economics* 8, 81–87.

Gujarati, D. (1995): *Basic Econometrics*, 3rd ed., McGraw-Hill, New York.

Hackethal, A.; Schmidt, R. (2005): Structural Change in the German Banking System?, Working Paper, University of Frankfurt.

Hall, A.D., Anderson, H.M. and Granger, C.W.J. (1992): A Cointegration Analysis of Treasury Bill Yields, *The Review of Economics and Statistics* 74, 116–126.

Hamburger, M.J. and Platt, E.N. (1976): The Expectations Hypothesis and the Efficiency of the Treasury Bill Market, *Review of Economics and Statistics* 57, 190–199.

Hannan, E.J. and Quinn, B.G. (1979): The Determination of the Order of an Autoregression, *Journal of the Royal Statistical Society*, B41, 109–195.

Hardouvelis, G. (1994): The Term Structure Spread and Future Changes in Long and Short Rates in the G7 Countries, *Journal of Monetary Economics* 33, 255–283.

Harris, R. and Sollis, R. (2005): *Applied Time Series Modelling and Forecasting*, Wiley, Chichester.

Hendry, D.F. and Anderson, G.J. (1977): Testing Dynamic Specification in Small Simultaneous Models: An Application to a Model of Building Society Behavior in the United Kingdom; in: Intrilligator, M.D. (ed.): *Frontiers of Quantitaive Economics*, Amsterdam, 361-383.

Hicks, J.R. (1946): *Value and Capital*, 2nd ed., Oxford.

Holden, D. and Perman, R. (2007): Unit Roots and Cointegration for the Economist, in Rao, B. (ed.): Cointegration for the Applied Economist, 2nd ed., Palgrave Macmillan, Hampshire, pp. 43–100.

Hördahl, P., Tristani, O. and Vestin, D. (2006): A Joint Econometric Model of Macroeconomic and Term-Structure Dynamics, *Journal of Econometrics* 131, 405–444.

Hsu, C., Kugler, P. (1997): The Revival of the Expectations Hypothesis of the US Term Structure of Interest Rates, *Economics Letters* 55, 115–120.

Hull, J. (2005): *Options, Futures and Other Derivatives*, 6th ed., Prentice Hall, Upper Saddle River, New Jersey.

Hull, J. (2007): *Risk Management and Financial Institutions*, Pearson Education, Upper Saddle River, New Jersey.

Ingersoll, J.E. (1987): *Theory of Financial Decision Making*, Rowman & Littlefield Publishers, Inc., Maryland.

Jarrow, R. and Turnbull, S. (2000): *Derivative Securities*, 2nd ed., South-Western College Publishing, Cincinnati.

Jarque, C.M. and Berra A.K. (1987): A Test for Normality of Observations and Regression Residuals *International Statistical Review* 55, pp. 163–172.

Israelsen, C.L. (2003): Sharpening the Sharpe Ratio, *Financial Planning* 33, 49–51.

Johansen, S. (1988): Statistical Analysis of Cointegrated Vectors, *Journal of Economic Dynamics and Control* 12, 231–254.

Johansen, S. (1991): Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models, *Econometrica* 59, 1551–1580.

Johansen, S.J. and Juselius, K. (1992): Testing Structural Hypothesis in a Multivariate Cointegration Analysis of the PPP and the UIP for the U.K., *Journal of Econometrics* 53, 211–244.

Jondeau, E. and Ricart, R. (1999): The Expectations Hypothesis of the Term Structure: Tests on US, German, French and UK Euro-rates, *Journal of International Money and Finance* 18, 725–750.

Jones, F.J. (1991): Yield Curve Strategies, *Journal of Fixed Income*, September, 43–51.

Jorion, P. (1992): Term Premiums and the Integration of Eurocurrency Markets, *Journal of International Money and Finance* 11, 17–39.

Jorion, P. (1997): *Value at Risk: The New Benchmark for Controlling Market Risk*, McGraw Hill, New York.

Jorion, P. and Mishkin, F. (1991): A Multicountry Comparison of Term-Structure Forecasts at Long Horizons, *Journal of Financial Economics* 29, 59–80.

Kahneman, D., and Tversky, A. (1979): Prospect Theory: An Analysis of Decision under Risk, *Econometrica* 47, 262–292.

Kaplan, P.D. and Knowles, J.A. (2004): Kappa: A Generalized Downside Risk-Adjusted Performance Measure, *Journal of Performance Measurement* 8, 42–54.

Kwiatkowski, D., Phillips, P.C., Schmidt, P. and Shin, Y. (1992): Testing the Null Hypothesis of Stationary Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?, *Journal of Econometrics* 54, 159–178.

Lanne, M. (2000): Near Unit Roots, Cointegration, and the Term Structure of Interest Rates, *Journal of Applied Econometrics* 15, 513–529.

Levy, H. (2006): *Stochastic Dominance: Investment Decision Making Under Uncertainty*, 2nd ed., Boston, MA: Springer Science+Business Media.

Lhabitant, F.S. (2004): *Hedge Funds: Quantitative Insights*, Wiley, West Sussex.

Littermann, R and J. Scheinkman (1991): Common Factors Affecting Bond Returns, *Journal of Fixed Income* 1, 54–61.

Longstaff, F.A. and Swartz, E.S. (1992): Interest Rate Volatility and the Term Structure: A Two Factor General Equilibrium Model, *Journal of Finance* 47, 1252–1282.

Lamsdaine, R. and Papell, D. (1997): Multiple Trend Breaks and the Unit Root Hypothesis, *Review of Economics and Statistics* 79, 212–218.

Mankiw, N. and Miron, J.A. (1986): The Changing Behavior of the Term Structure of Interest Rates, *The Quarterly Journal of Economics* 101, 211–228.

Mankiw, N. and Summers, L.H. (1984): Do Long Term Rates Overreact to Short-Term Interest Rates?, *Brookings Papers on Economic Activity* 1, 223–242.

Mann, S.V. and P. Ramanlal (1997): The Relative Performance of the Yield Curve Trading Strategies, *Journal of Portfolio Management* 23, 64–70.

Marcowitz, H.M. (1952): Portfolio Selection, *Journal of Finance* 7, 77–91.

Martellini, L., Priaulet, P. and S. Priaulet (2002): Understanding the Butterfly Strategy, *Journal of Bond Trading and Management* 1, 9–19.

Martellini, L., Priaulet, P. and S. Priaulet (2003): *Fixed-Income Securities: Valuation, Risk Management and Portfolio Strategies*, Wiley, West Sussex, England.

McCulloch, J.H. (1971): Measuring the Term Structure of Interest Rates, *Journal of Business* 44, 19–31.

Mercer, J., Moore, M.E. and D.B. Winters (2009): Do Traders Benefit from Riding the T-Bill Yield Curve? *Journal of Portfolio Management* 26, 131–140.

Mishkin, F. (1992): *The Economics of Money, Banking and Financial Markets*, 3rd ed., Harper Collins Publishers, New York.

Modigliani, F. and Modigliani, L. (1997): Risk-Adjusted Performance, *Journal of Portfolio Management* 23, 45–54.

Modigliani, F. and Sutch, R. (1966): Innovations in Interest Rate Policy, *American Economic Review* 56, 178–197.

Mustafa, M. and Rahman M. (1995): Cointegration between US Short-Term and Long-Term Interest Rates (Both Nominal and Real), *Applied Financial Economics* 5, 323–327.

Nawrocki, D. (1991): Optimal Algorithms and Lower Partial Moments: Ex-Post Results, *Applied Economics* 23, 465–470.

Nawrocki, D. (1999): A brief History of Downside Risk Measures, *Journal of Investing* 8, 9–26.

Nelson, C.R. and Plosser, C.I. (1982): Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications, *Journal of Monetary Economics* 10, 139–162.

Nelson, C.R and Siegel, A.F. (1987): Parsimonious Modelling of Yield Curves, *Journal of Business* 60, 473–489.

Newey, W.K. and West, K.D. (1987): A Simple Positive-Definite Heteroscedasticity and Autocorrelation-Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Nourzad, F. and Grennier, S. (1995): Cointegration Analysis of the Expectations Theory of the Term Structure, *Journal of Economics and Business* 47, 281–292.

Osteryoung, J.S., McCarty, D.E. and Roberts, G. (1981): Riding the Yield Curve with Treasury Bills, *The Financial Review* 16, 57–66.

Pelaez, R. (1997): Riding the Yield Curve: Term Premiums and Excess Returns, *Review of Financial Economics* 6, 113–119.

Patterson, K. (2010): *A Primer for Unit Root Testing*, Palgrave Macmillan, Hampshire.

Perron, P. (1989): The Great Crash, the Oil Price Shock and the Unit Root Hypothesis, *Econometrica* 57, 1361–1401.

Phillips, P.C. and Perron, P. (1988): Testing for a Unit Root in Time Series Regression, *Biometrika* 75, 335–346.

Pindyck, R.S. and Rubinfeld, D.L. (1998): *Econometric Models and Econometric Forecasts*, 4th ed., McGraw-Hill, Boston.

Quandt, R. (1960): Test of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes, *Journal of the American Statistical Association* 55, 324–330.

Reichling, P., Afanasenko, D. (2010): The German Banking System: Regulation, Structure and Basel II Implementation; in: Yepifanov, A.; Shkolnik, I. (ed.): Basel II: Problems and Prospects of Usage in National Banking Systems, Sumy, 12–37.

Rockafellar, R.T and Uryasev, S. (2002): Conditional Value at Risk for General Loss Distributions, *Journal of Banking and Finance* 26, 1443–1471.

Sargan, J.D. (1984): Wages and Prices in the United Kingdom: A Study in Econometric Methodology, in Hendry, D.F. and Wallis, K.F.: Econometrics and Quantitative Economics, Blackwell, Oxford.

Scholz, H. (2007): Refinements to the Sharpe Ratio: Comparing Alternatives for Bear Markets, *Journal of Asset Management* 7, 347–357.

Schwarz, G. (1978): Estimating the Dimension of a Model, *Annals of Statistics* 6, 461–464.

Sharpe, W. (1966): Mutual Fund Performance, *Journal of Business* 39, 118–138.

Shea, G.S. (1992): Benchmarking the Expectations Hypothesis of the Term Structure: an Analysis of Cointegrating Vectors, *Journal of Business and Economics Statistics* 13, 347–365.

Shiller, R.J., Campbell, J.Y. and Schoenholtz, K.L. (1983): Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates, *Brookings Papers on Economic Activity* 1, 173–217.

Sorensen, P.B., Whitta-Jacobsen, H.J. (2005): *Introducing Advanced Macroeconomics: Growth and Business Cycles*, McGrow-Hill, Berkshire.

Sortino, F., Kordonsky, M. and Forsey, H. (2006): Evidence-Based Portfolio Management, Working Paper.

Sortino, F. and Price, L. (1994): Performance Measurement in a Downside Risk Framework, *Journal of Investing*, Fall, 1519–1533.

Sortino, F., van der Meer, R., Platinga, A. (1999): The Dutch Triangle, *Journal of Portfolio Management* 26, 50–58.

Spearman, C. (1904): The Proof and Measurement of Association Between Two Things, *American Journal of Psychology* 15, 72–101.

Stock, J.H. (1987): Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors, *Econometrica* 55, 1035–1056.

Stock, J.H. and Watson, M. (1988): Testing for Common Trends, *Journal of American Statistical Association* 83, 1097–1107.

Svensson, L. (1994): Estimating and Interpreting Forward Interest Rates: Sweden 1992 – 94, CEPR Discussion Paper 1051.

Szegö, G. (2002): Measures of Risk, *Journal of Banking and Finance* 26, 1253–1272.

Tabak, B. (2009): Testing the Expectations Hypothesis in the Brazilian Term Structure of Interest Rates: A Cointegration Analysis, *Applied Economics* 41, 2681–2689.

Tasche, D. (2002): Expected Shortfall and Beyond, *Journal of Banking and Finance* 26, 1519–1533.

Taylor, M. (1992): Modelling the Yield Curve, *The Economic Journal* 102, 524–537.

Thomas, R.L. (1997): *Modern Econometrics: An Introduction*, Addison-Wesley, Harlow.

Van Horne, J.C. (1974): *Financial Management and Policy*, 4th ed., Prentice Hall, New Jersey.

Vasicek, O.A. (1977): An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics* 5, 177–188.

Vasicek, O.A. and Fong, H.G. (1982): Term Structure Modeling Using Exponential Splines, *Journal of Finance* 37, 339–348.

Walsh, C.E. (2003): *Monetary Theory and Policy*, 2nd ed., The MIT Press, Cambridge.

Wu, T. (2001): Monetary Policy and the Slope Factor in Empirical Term Structure Estimation, Federal Reserve Bank of San Francisco, Working Paper, 2002-07.

Wu, J.L. and Chen, S.L. (2001): Mean Reversion of Interest Rates in the Eurocurrency Market, *Oxford Bulletin of Economics and Statistics* 63, 459–473.

Zhang, H. (1993): Treasury Yield Curves and Cointegration, *Applied Economics* 25, 361–367.

Zivot, E. and Andrews, D. (1992): Further Evidence of the Great Crash, the Oil Price Shock and the Unit Root Hypothesis, *Journal of Business and Economic Statistics* 10, 251–270.