Maastricht University

The term structure of interest rates : a panel data analysis

Citation for published version (APA):

Bams, W. F. M. (1999). *The term structure of interest rates : a panel data analysis*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. https://doi.org/10.26481/dis.19990326wb

Document status and date:
Published: 01/01/1999

DOI:
10.26481/dis.19990326wb

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 20 Apr. 2024

The Term Structure of Interest Rates: A Panel Data Analysis

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. dr. A.C. Nieuwenhuijzen Kruseman,
volgens het besluit van het College van Decanen, in het openbaar te verdedigen
op vrijdag 26 maart 1999 om 12.00 uur

door

Wilhelmus Fransiscus Maria Bams

Promotores :

Prof. dr. P.C. Schotman
Prof. dr. C.C.P. Wolff

Beoordelingscommissie :

Prof. dr. C.G. Koedijk	(Voorzitter)
Prof. dr. Th.E. Nijman	(Katholieke Universiteit Tilburg)
Prof. dr. G.A. Pfann	(Universiteit Maastricht)
Prof. dr. R. Stapleton	(Strathclyde University)

Financiële steun :

NWO 510-31-210

Preface

This dissertation is the result of four years of work at the Finance Department of Maastricht University, under the auspices of the Dutch Foundation for Scientific Research's division of Longitudinal Econometrics. I first experienced the challenge of working on a dissertation when I assisted Cees Dert, my Master's Thesis supervisor, with a project for his dissertation. It was at that time that I decided I would enjoy launching such a project myself. By happy coincidence I became acquainted with Peter Schotman, who informed me about an opening in the Department. After a more formal discussion with him and Christian Wolff, I was accepted to work on this project.

Looking back at the past four years, it is clear that writing a dissertation involves much more than a narrow focus on one specific subject. As a PhD student at the Finance Department of Maastricht University, I was surrounded by a group of people who were actively involved and prolific writers in the field of finance. I enjoyed our many discussions on the latest hot topics. Although a university is often thought of as a place for experts, the combination of their expertise provides the broad education which one seeks. I benefitted a great deal from my colleagues expertise.

Although writing this dissertation involved spending every free minute thinking ahead, programming algorithms, and writing down results; I also had the opportunity to attend many conferences and to spend time at the Ecole des Hautes Etudes en Sciences Sociales in Marseille and the Chicago Graduate School of Business. Many thanks go to my supervisors for enabling me to take the interesting opportunities to meet new people working on similar projects.

Having spent the first eighteen years of my life in Maastricht, I was happy to return to this beautiful old city with its many good pubs and restaurants. And with the University only minutes away from the city centre, I visited them frequently. I was also happy to be close to my family again, making it easy to invite myself for dinner at my mother's on a regular basis.

Even though this dissertation is a solo performance - there is a whole stage crew without whom it could not have been completed. I am firstly indebted to the Dutch Foundation for Scientific Research (NWO) for their financial support of both my research and my attendance at conferences and other universities.

My largest debt by far is to my daily supervisor, Peter Schotman, who showed me how exciting the world of research can be. The vast improvement in my econometric and

research skills can be accredited to him, and every chapter of this dissertation has been polished by his touch. Peter encouraged me to present my papers frequently and to join him at international conferences, where I acquired his habit of carrying a notebook to jot down derivations. But I especially appreciated our good relationship- I greatly enjoyed our frequent conversations over a glass of beer.

A special word of thanks also goes to my other supervisor, Christian Wolff. Despite his busy work as Dean of the Economics Faculty during the first two years of my research, Christian still found the time to keep both myself and my research evolving along the right path. I enjoyed conducting one research project directly with him, and greatly benefitted from his economic intuition. I thank Christian for the personal interest he took in me and the frequent assistance he gave, but still intend to beat him at golf one of these days!

Ronald Mahieu had almost completed his dissertation under the same supervisors at the time that I first started at the Finance Department, and was therefore a great help with the many practical and theoretical questions that I had. Ronald also became a good friend and I typically celebrated Carnival and New Year's Eve with him and Marian. I hope that we will have many future projects together.

Another instant good friend at Maastricht University was Arjan van Bussel. Since Arjan was working on a project closely related to mine, we had almost daily discussions about our progress and the problems we faced. Moreover, because he spent as many late nights and weekends at the department as I did, he kept me company and joined me in many more relaxing activities like dinner or a beer (or two). Thanks!

Although writing a dissertation did turn out to be much different from my expectations four years ago, I enjoyed every part of it and would immediately do it all over again if given the chance.

Dennis Bams
December, 1998.

Contents

1 An Introduction to the Class of Affine Term Structure Models **1**

1.1 Introduction 1
1.2 Why is term structure research important 2
1.3 Factors affecting fixed income securities 2
1.4 The term structure of interest rates 3
1.5 How to model the term structure of interest rates 4
1.6 Affine term structure models 6
1.7 Endogenous no-arbitrage models 7
1.8 Exogenous no-arbitrage models 11
1.9 Endogenous equilibrium models 13
1.10 Solutions to the Ricatti equations 16
1.11 Research objectives 20

2 Curve Fitting Methods **23**

2.1 Introduction 23
2.2 Econometric framework 25
2.3 Approximate function specification 28
2.4 Data 31
2.5 Results 34
2.6 Concluding remarks 54

3 The Time Series and Cross Sectional Properties of Interest Rates Risk Premia **55**

3.1 Introduction 55
3.2 Background and notation 57
3.3 The econometric framework 59
3.4 Data 64
3.5 Estimation results 70
3.6 Concluding Remarks 77

4 Direct Estimation of the Risk Neutral Factor Dynamics of Affine Term Structure Models 81
4.1 Introduction . . . 81
4.2 Affine term structure models . . . 84
4.3 Gaussian models . . . 87
4.4 Why pooling? . . . 89
4.5 Econometric model . . . 92
4.5.1 Specification . . . 93
4.5.2 Estimation and Testing . . . 96
4.6 Data . . . 98
4.7 Empirical results . . . 99
4.8 Conclusions . . . 111

5 An Empirical Investigation to the Longstaff-Schwartz Term Structure Model 113
5.1 Introduction . . . 113
5.2 Model specification . . . 115
5.3 Econometric specification . . . 120
5.3.1 Time series approach . . . 121
5.3.2 Panel data approach . . . 122
5.4 Data . . . 125
5.5 Results . . . 127
5.5.1 CIR one-factor panel data model . . . 128
5.5.2 LS time series model . . . 132
5.5.3 LS panel data model . . . 137
5.5.4 Cap pricing . . . 142
5.6 Concluding remarks . . . 144

6 Summary and Concluding Remarks 147
6.1 Introduction . . . 147
6.2 Summary . . . 148
6.3 Directions for further research . . . 151

Nederlandse samenvatting / Dutch summary 153

Bibliography 157

Curriculum Vitae 163

List of Tables

2.1 Bond Prices . . . 33
2.2 Summary statistics : Spline model (spread weighted errors) . . . 39
2.3 Summary statistics: Spline model (duration weighted errors) . . . 40
2.4 Summary statistics: Nelson-Siegel method (duration weighted errors) . . . 41
2.5 Summary statistics: Extended Nelson-Siegel method . . . 42
2.6 Evolution of forward rate in spline models . . . 43
2.7 Evolution of forward rate in Nelson-Siegel models . . . 44
2.8 Error term analysis . . . 53

3.1 Summary statistics . . . 68
3.2 Correlation matrix . . . 69
3.3 Results univariate regressions . . . 71
3.4 Results pooled regression . . . 72
3.5 Results fixed maturity effects: $\psi(\tau)$ . . . 73
3.6 Results fixed maturity/random time effects panel model . . . 75
3.7 Results risk premium specification: $Z(\tau)$ . . . 76

4.1 Monte Carlo Results Vasicek Yield Curves . . . 91
4.2 Parameter Estimates . . . 105

5.1 Summary statistics . . . 126
5.2 Correlation matrix . . . 127
5.3 Estimation Results CIR Model . . . 128
5.4 Summary Statistics Residuals . . . 131
5.5 Factor Loading . . . 132
5.6 Estimation Results LS Model, Time Series Approach . . . 133
5.7 Factor Loadings . . . 136
5.8 Estimation Results LS Panel Model . . . 137
5.9 Factor Loadings . . . 139
5.10 Summary Statistics Residuals . . . 141

List of Figures

2.1 Term Structures Cubic Spline Model 37
2.2 Term structures Nelson-Siegel Model 38
2.3 Time series of Interest rates Confidence Intervals 45
2.4 Time series of Interest rates Confidence Intervals 46
2.5 Fit Spread Weighted Cubic Spline Method 47
2.6 Fit Duration Weighted Cubic Spline Model 48
2.7 Fit Duration Weighted Nelson-Siegel Model 49
2.8 Fit Duration Weighted Extended Nelson-Siegel Model 50

3.1 US Interest Rates 66
3.2 Volatilities error term 74
3.3 Risk premium specification 77

4.1 Yield Curve Pooling 90
4.2 Yield curve fitting 94
4.3 US Interest Rates 100
4.4 Data Summary 101
4.5 A second factor 102
4.6 Discount yields and forward rates 103
4.7 Average Yield Curve 106
4.8 Term Structure of Volatilities 107
4.9 Implied spot rates 108

5.1 US Interest Rates 126
5.2 Implied Factors 129
5.3 Fit Term Structure 130
5.4 GARCH Volatility 134
5.5 Implied Term Structure 135
5.6 Fit Term Structure 138
5.7 Implied Factors 140
5.8 Cap Rates 143
5.9 Cap Maturities 143

Chapter 1

An Introduction to the Class of Affine Term Structure Models

1.1 Introduction

This dissertation deals with affine models for the term structure of interest rates. The term structure of interest rates measures the relationship among the yields on default-free securities that differ only in their term to maturity. Affine models are a particular class of models to describe the term structure of interest rates. The aim of this chapter is to give an introduction to this class of term structure models. The term structure determines the discount rates that are used to price fixed income securities. Fixed income securities as the name already suggests deals with investments that are bought primarily to earn steady income from regular interest payments in the form of coupons.

The basic example of a fixed income instrument is a coupon bearing government bond. This is an instrument that pays a steady flow of coupons at prespecified days in the future, and repays the principal at maturity of the bond. The current value of the bond is the total discounted value of all future payments that are related to this bond. These payments are discounted against the current term structure of interest rates and hence the term structure is an important ingredient for fixed income pricing.

In the following, we start with a general discussion on the term structure of interest rates. We consider issues that deal with the importance of term structure research, which factors affect fixed income securities and why the term structure of interest rates is an important concept to model fixed income securities. Then we outline what different classifications exist to model term structures. Affine term structure models are an important subclass that is motivated and described next. We show the implications for endogenous and exogenous term structure models when we deal with affine models. We end with an outline of the research objectives of this dissertation.

1.2 Why is term structure research important

Fixed income securities are important to institutional investors, like banks, insurance companies and mutual funds, because their investment portfolio consists for an important part of fixed income securities. Since these institutions are usually acting on behalf of individuals who in one way or the other have supplied these institutions with funds, the term structure of interest rates is also important to individuals. Other situations in which the term structure becomes important for individuals is for example when they have to agree upon a mortgage loan or when they have to start saving for their retirement. Relevant issues in those cases are for example the choice between a fixed rate and an adjustable-rate mortgage, and the choice between a bond fund and a money market fund. The term structure of interest rate is the key determinant in those cases.

Research of the term structure of interest rates is important for several reasons. Melino (1988) states that economists have had a long interest in the term structure, which has been motivated by a desire to understand the link between monetary policy and real economic activity. According to the conventional wisdom, the central banks mainly affect short-term interest rates such as the yields on Treasury bills. Real economic activity, on the other hand, is more closely linked to the yields on bonds with the same maturity as physical capital, say in the order of ten to twenty years. According to this view, it is crucial that we understand the factors which influence the relative yields on these different types of securities, in order to understand the impact of central bank actions on the real side of the economy.

Estrella and Hardouvelis (1991) state that the importance of the term structure lies in the fact that it embodies the market's anticipation of future events. Therefore an explanation of the term structure gives us a way to extract this information and to predict how changes in the yield curve will affect real economic activity.

Finally, the term structure of interest is used as an instrument for the pricing of fixed income securities. The most important determinant is the interest rate. Since the fixed income market is extremely large regardless of whether size is measured by quantities outstanding, or quantities trading, it is of interest that we understand how fixed income instruments are priced, hence that we understand the term structure of interest rate. For hedging purposes it is interesting to gain insight in the dynamics of the term structure, since it determines how the value of a fixed income portfolio changes.

1.3 Factors affecting fixed income securities

So far we have dealt with fixed income securities that are fully determined by the periodical coupon payments, which are certain, and by the time to maturity of the bond, after which the principal is repaid. In practice the same definition is also used for securities that are similar to fixed income securities. Examples are bonds that differ in that they can include option features, like callable bonds or convertible bonds. Other features that influence the

price of a bond is first its credit rating, which reflects the probability that the borrower does not fully pay the future coupons. Second, the liquidity of the bond reflects the trade in the bond and hence determines whether a bond can actually be sold against the price that is quoted on the market. Third, in general US bonds are subject to different kinds of taxes, which influence their prices. Fourth, there exist bonds for which the size of the coupon payments is uncertain, like structured notes.

It should be noted that even in the case where future payments are certain, changing economic conditions can influence the value of a bond. The most important determinant of the price of a bond is the level of interest rates. In this dissertation we will only consider government bonds, bills and notes since these are typically the 'cleanest' type of fixed income instruments. US government securities carry the highest possible credit rating, since they are fully backed by the government. They are thus the safest type of investment available for a guaranteed payment of principal and coupon. This type of securities does not include any option features, which makes it easier to determine their price. Publicly traded US government securities are more liquid and marketable than any other sort of portfolio investment.

An important part of this dissertation deals with pricing of fixed income securities. We will estimate models using yield curve data. The models are derived under the assumption of no-arbitrage, which makes them applicable to price all kinds of other interest rate derivative securities as well. For example, even though we do not include bonds in our data set that include all sorts of option features, our framework does allow us to price these option features.

1.4 The term structure of interest rates

A default free coupon bearing bond is specified in terms of the level of the periodical coupon payment and in terms of the time to maturity of the bond. It is possible that many different bonds exist in the market that differ with respect to these two features. Fixed income research inquires what market forces are responsible for the prices and changes in the prices of bonds. The fact that fixed income instruments differ with respect to coupons, coupon payment dates and maturities results in a bond market that is complex and not very transparent. In order to simplify the analysis we introduce the term structure of interest rates, which enables the simultaneous treatment of all different fixed income instruments. Once the term structure of interest rates is known, we can price all bonds.

The term structure of interest rates is helpful in determining the current value of some future payment. More specifically, it describes the relation between the time of a future cashflow and the interest rate that is applicable to obtain the current value of the future cashflow. In general, cashflows that occur at different times in the future are subject to different discount rates. Since all bond can be viewed as instruments that guarantee certain future payments at different dates, their price is the total discounted value of

all these future payments, where discounting occurs against the current prevailing term structure of interest rates. Hence, the level and changes in the term structure of interest rates are informative on the prices and on the changes in the prices of bonds.

The term structure of interest rates is usually increasing, which means that longer term cashflows are discounted at a higher annual rate than short term cashflows. Occasionally, they are inverted, which implies that the market requires a higher annual yield for money that is held for shorter time. Sometimes term structures are humped shaped.

1.5 How to model the term structure of interest rates

Although yields with different maturities can be treated as separate variables, it makes sense to study them simultaneously. Typically yields are influenced by the same variables, only the degree in which they are influenced is different. An analysis of the yield curve is particular challenging because instead of dealing with one variable or a couple of variables, an entire curve has to be modeled. Melino (1988) states that the principal concern in the mainstream economic literature has been with the pricing of bonds identical in every aspect except for maturity. In particular, economists have studied the pricing of pure discount bonds, that are not only free of default risk, but also free of call or other options. The emphasis on this very special aspect of bond pricing is important since it is one of the main building blocks of fixed income securities.

In a world of certainty equilibrium forward rates must coincide with future spot rates, but when uncertainty about future rates is introduced the analysis becomes much more complex. Previous theories of the term structure have taken the certainty model as their starting point and have proceeded by examining stochastic generalizations of the certainty equilibrium relationship. All these early theories belong to four different strands of thoughts.

First, there is the pure expectations hypothesis in all its different forms, which has the following implications: (a) the forward rate is equal to the expected future spot rate; (b) the return on holding a long term discount bond to maturity is equal to the expected return on repeated investments in a series of one period discount bonds; (c) the expected holding period return is the same for discount bonds of all maturities.

Second, the liquidity preference theory, advanced by Hicks (1946), asserts that because of risk averse market participants, forward rates will be systematically larger than expected future spot rates. The difference between the two is called the term premium and this premium is usually increasing with maturity to induce investors to hold longer term securities.

A third explanation of the term premium is the market segmentation theory of Culbertson (1957). This theory states that investors have strong maturity preferences and that bonds of different maturities trade in separate and distinct markets. When demand and supply for a bond of a certain maturity is little affected by the price of a bond with

neighboring maturity, then there is no reason for the term premium to be positive and to be increasing with maturity as in the liquidity preference theory.

Fourth, in the preferred habitat theory Modigliani and Sutch (1966) intend their approach as a plausible rationale for term premiums which does not restrict them in sign or monotonicity.

The previous theories basically use the time series information that is available in yield curve data. The different theories focus on the determinants of the term premium. Explaining and testing term structure models in this framework is interesting as no assumptions are made on the dynamics of the relevant variables. The approach is however not applicable for derivatives pricing. For that purpose typically cross sectional relationships are important.

In the modern term structure literature a different approach is pursued. The entire yield curve is modeled by imposing that there exists a limited number of factors that determine the yield curve. All yields on the curve are determined by the dynamics of the underlying factors and the relation only differs with maturity. DeMunnik (1992) provides a classification of different term structure approaches. He defines the class of models that are determined by some underlying factor process as the Indirect Approach, as opposed to the Direct Approach in which the process of a bond is explicitly formulized. A severe disadvantage of the direct approach is the fact that this process has to capture specific features of the bond, which has been proven difficult.

The indirect approach starts with the specification of one of more factors processes on which all other interest rate derivative securities are depending. In the so-called single factor model, the instantaneous spot rate is adopted as the factor that is driving the dynamics of the yield curve. A further classification of term structure models is related to whether the currently observed term structure of interest rates follows from the underlying factor processes as well, or whether it is assumed to be known and explicitly incorporated in the model. In the first case the possible shape of this term structure is endogenously implied by the stochastic characteristics of the underlying factors. In the second case the parameters of the underlying factor processes are time dependent to ensure the currently observed term structure is modeled without error. We refer to this second class as exogenous term structure models.

There exist mainly two approaches to derive yield curve implications from the underlying factor processes. The first approach is the no-arbitrage approach that prices all securities in an economy in such a way that arbitrage opportunities are precluded. The second approach is the equilibrium approach in which an economic framework is set up and where the prices follow after utility maximization by some representative agent. Often both approaches lead to the same yield curve representations, which means that no-arbitrage models often have an equilibrium interpretation as well.

The indirect approach, in which a number of factors are adopted has a number of advantages. First, it sets out a complete framework for all interest rate derivative securities, that are priced such that arbitrage opportunities between the different securities are

precluded. Second, since the same factors determine the prices of different securities, the indirect approach provides a framework for relative pricing. It is able to describe how the prices of different securities are changed relative to each other, because the same factors determine the changes in the prices of the securities. The indirect approach makes it easier to model the dynamics of the entire yield curve, since it reduces the problem to model the dynamics of a couple of factors.

1.6 Affine term structure models

To model yield curves, we restrict ourselves to the class of affine term structure models. In this class the factors have drift and diffusion specifications that are affine in the level of the factors and also the yields are affine functions of the factors. Special cases of the affine class of term structure models include Chen (1995), Chen and Scott (1992), Cox, Ingersoll and Ross (1985b), Duffie and Kan (1996), El Karoui and Lacoste (1992), El Karoui and Rochet (1989), Jamshidian (1989, 1990, 1991, 1992), Frachot, Janci and Lacoste (1992), Frachot and Lesne (1993), Heston (1991), Langetieg (1980), Longstaff and Schwartz (1992), Pennachi (1991) and Vasicek (1977).

Because of the affine relationship, this class of models is particular attractive for empirical work. Empirical research of affine term structure models include Brown and Dybvig (1985), Brown and Schaefer (1993), Chen and Scott (1992, 1993), DeJong (1997), DeMunnik and Schotman (1994), Duffie and Singleton (1995), Frachot, Janci and Lacoste (1992), Frachot and Lesne (1993), Heston (1991), Longstaff and Schwartz (1992a, 1992b, 1993a, 1993b), Pearson and Sun (1994), Pennachi (1991), Schotman (1996) and Stambaugh (1988).

We present three different approaches to introduce affine term structure models, the no-arbitrage approach as proposed by Duffie and Kan (1997) for endogenous term structure models, the no-arbitrage approach as proposed by Frachot and Lesne (1993) for exogenous term structure models and the equilibrium approach which originates from Cox, Ingersoll and Ross (1985a).

The equilibrium approach deals with endogenous term structure models and results in an equation that relates the price of an asset to the values of some underlying factors. It expresses the price of a discount bond as a function of the dynamics of the underlying factors. No-arbitrage models are based upon the fact that in complete markets arbitrage opportunities do not exist. The exogenous no-arbitrage approach encompasses the models of the Heath, Jarrow and Morton (1992, HJM) type and includes the endogenous no-arbitrage models as a special case.

In the following these approaches are described. All three approaches lead to Ricatti equations that express the coefficients of the discount bond prices and the yields in terms of the factor process parameters under the risk adjusted probability measure. After deriving the Ricatti equations, we take a closer look at particular cases of affine class models that

lead to closed form solutions of the Ricatti equations. Also we briefly deal with the most important properties of affine class models.

1.7 Endogenous no-arbitrage models

Duffie and Kan (1996) define and analyze a multifactor affine term structure model. The model is affine in the sense that all yields are an affine function of the underlying factors. The factor process is described by a Markovian diffusion process in which both the drift part and the variance part are affine in the factor levels.[1] The model specifies simple relationships among yields and provides term structure derivative prices, consistent with the absence of arbitrage. Moreover the model is computationally tractable and convenient for empirical work.

In no-arbitrage models a process for the underlying factors is set up and next the process is solved for the term structure. The underlying principle is to determine the drift specification of discount bonds in terms of the factor dynamics. Because of the absence of arbitrage, the drift of any traded derivative security is, under the risk adjusted probability measure, equal to the risk free rate. A general result for term structure models is that if the factors are affine then discount bonds are exponential affine and yields at any given maturity τ turn out to be a τ-dependent affine function of the underlying factors.

Suppose X_t represents a $(K \times 1)$ vector of factors that drive the dynamics of the yield curve

$$X_t = \begin{bmatrix} X_{1t} \\ \vdots \\ X_{Kt} \end{bmatrix} \tag{1.1}$$

where we introduce an affine diffusion process for the factors

$$dX_t = [CX_t + d]\,dt + v(X_t)\,dW_t \tag{1.2}$$

with C a $(K \times K)$ matrix and d a $(K \times 1)$ vector of parameters, and W_t denotes Brownian motion under the actual probability measure. The volatility part is given by the $(K \times K)$ matrix $v(X_t)$ that determines the instantaneous factor covariance matrix

$$V(X_t) = v(X_t)\,v(X_t)' = \begin{pmatrix} \alpha_{11} + \beta_{11}X_t & \cdots & \alpha_{1K} + \beta_{1K}X_t \\ \vdots & & \vdots \\ \alpha_{K1} + \beta_{K1}X_t & \cdots & \alpha_{KK} + \beta_{KK}X_t \end{pmatrix} \tag{1.3}$$

where α_{ij} are scalars and β_{ij} are $(1 \times K)$ vectors of parameters, $\forall i, j$. The instantaneous factor variances are affine functions of the factors. Since X_t is stochastic, the vectors β_{ij} generate stochastic volatility unless they are all zero. In the latter case the factors

[1] In the following we mean by an affine factor process, a diffusion process in which both drift and variance are affine functions of the factor levels.

follow a Gaussian Markov process. This is also referred to as the constant volatility or the Gaussian case. With π we denote the parameters describing the factor process under the actual probability measure, where

$$\pi = \{C,\ d,\ \alpha_{11},\ \ldots,\ \alpha_{KK},\ \beta_{11},\ \ldots,\ \beta_{KK}\} \tag{1.4}$$

The factors are generally not observable and following Pearson and Sun (1994) and Chen and Scott (1993), Duffie and Kan (1996) propose a change of basis to observable yields. Alternative methods to deal with the unobserved factors is to approximate them with observable variables (for example use the 1 month T-bill rate as an approximation of the instantaneous spot rate) or to treat them as latent variables, incorporating a time series process for the factors.

In order to price bonds and other term structure derivative securities an assumption about the market prices of risk have to be made. Each factor that is priced has a market price of risk that differs from zero. Duffie and Kan (1996) assume that the market price of risk for factor i is proportional to the variance in the factor diffusion process. The risk adjusted Brownian motion, W_t^*, therefore is defined as

$$dW_t^* = dW_t + v(X_t)^{-1} \Lambda tr[V(X_t)]\, dt \tag{1.5}$$

where Λ denotes a $(K \times K)$ diagonal matrix which has as i-th diagonal element the market price of risk associated with the i-th factor. The factor diffusion process under the risk adjusted probability measure follows immediately as

$$dX_t = [C^* X_t + d^*]\, dt + v(X_t)\, dW_t^* \tag{1.6}$$

where

$$C^* = C - \Lambda \begin{pmatrix} \beta_{11} \\ \vdots \\ \beta_{KK} \end{pmatrix} \tag{1.7}$$

and

$$d^* = d - \Lambda \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{KK} \end{pmatrix} \tag{1.8}$$

Notice that only parameters in the drift part of the factor process change, when we switch to the risk adjusted probability measure, the volatility part remains the same. We summarize the factor parameters under the risk adjusted probability measure in the set

$$\pi^* = \{C^*,\ d^*,\ \alpha_{11},\ \ldots,\ \alpha_{KK},\ \beta_{11},\ \ldots,\ \beta_{KK}\} \tag{1.9}$$

With $Y(t,T)$ we denote the yield at time t for cashflows with maturity date T, and with $P(t,T)$ we denote the price of a discount bond at time t that matures at time T. We adopt an exponential affine relationship between the discount bond prices and the factors

$$P(t,T) = \exp[A(t,T) + B(t,T) X_t] \tag{1.10}$$

The $(1 \times K)$ vector of factor loadings is given by

$$B(t,T) = \begin{bmatrix} B_1(t,T) & \cdots & B_K(t,T) \end{bmatrix} \tag{1.11}$$

The i-th element of $B(t,T)$ is the factor loading for the i-th factor. At maturity, $t = T$, the principal is repaid and hence the price of a discount bond always equals unity. The following boundary conditions hold

$$A(T,T) = 0 \tag{1.12}$$

and

$$B(T,T) = o' \tag{1.13}$$

where o is a $(K \times 1)$ vector of zeros. The affine relationship between the yield curve and the underlying factors is put forward by

$$Y(t,T) = -\frac{\ln P(t,T)}{T-t} = -\frac{A(t,T)}{T-t} - \frac{B(t,T)}{T-t} X_t \tag{1.14}$$

The elements of $A(t,T)$ and $B(t,T)$ are specific functions of the parameters under the risk adjusted probability measure, π^*. Because the factors contain an arbitrary scale factor, the yields are invariant under scale transformations of the factors. Without loss of generality Duffie and Kan (1996) impose boundary conditions on the parameters of the factor processes such that for the derivatives of $A(t,T)$ and $B(t,T)$ at maturity it holds that

$$\lim_{t \to T} \frac{A(t,T)}{T-t} = 0 \tag{1.15}$$

and

$$\lim_{t \to T} \left(\frac{B(t,T)}{T-t} \right) = -\iota' \tag{1.16}$$

where ι is a $(K \times 1)$ vector of ones. From equation (1.14) it follows that these conditions restrict the intercept and the factor loadings at the very short end of the yield curve. For the instantaneous spot rate, r_t, it follows that it is the sum of the underlying factors

$$r_t = \lim_{T \to t} Y(t,T) = \iota' X_t \tag{1.17}$$

This also implies that under the given normalization for single factor affine term structure models, $K = 1$, the underlying factor is necessarily the instantaneous spot rate, $X_t = r_t$.

In the following we show that an affine factor process is consistent with an affine factor model for the yield curve. Moreover, we show that the parameters $A(t,T)$ and $B(t,T)$ depend in a specific way upon the parameters of the factor process under the risk adjusted probability measure. To demonstrate this, we derive the diffusion process of the discount bond price by application of Itô's lemma. Because of the no-arbitrage principle the drift part is, under the risk adjusted probability measure, equal to the risk free rate of return on the discount bond. Let $D(X_t,t)$ denote the drift part of the discount bond price, then it holds that

$$D(X_t,t) = r_t P(t,T) \tag{1.18}$$

Using the fact that the spot rate is equal to the sum of the factors, it follows that

$$D(X_t,t) = P(t,T)[\iota' X_t] \tag{1.19}$$

Applying Itô's lemma under the assumption that the price of a discount bond depends upon the factor processes, we also find the drift component in an alternative way.

$$D(X_t,t) = \frac{\partial P(t,T)}{\partial t} + \left(\frac{\partial P(t,T)}{\partial X_t}\right)[C^* X_t + d^*] + \frac{1}{2} tr\left(V(X_t)\frac{\partial^2 P(t,T)}{\partial X_t \partial X_t'}\right) \tag{1.20}$$

Substitute the appropriate derivatives and use the affine specifications for both the factor drift and the factor covariance specification to obtain

$$\begin{aligned} D(X_t,t) &= P(t,T)\left[\frac{dA(t,T)}{dt} + \frac{dB(t,T)}{dt}X_t\right] + P(t,T)B(t,T)[C^* X_t + d^*] + \\ &\quad + \frac{1}{2}P(t,T)\sum_{i=1}^{K}\sum_{j=1}^{K} B_i(t,T)B_j(t,T)(\alpha_{ij} + \beta_{ij}X_{it}) \end{aligned} \tag{1.21}$$

Under the adjusted probability measure the expressions in equations (1.19) and (1.21) are equal, hence

$$\begin{aligned} &\frac{dA(t,T)}{dt} + \frac{dB(t,T)}{dt}X_t + B(t,T)[C^* X_t + d^*] \\ &+ \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} B_i(t,T)B_j(t,T)(\alpha_{ij} + \beta_{ij}X_{it}) \\ = \; &\iota' X_t \end{aligned} \tag{1.22}$$

Since this equality holds for all possible maturities and for all possible factor values this results in the following Ricatti equations that establish the link between the parameters under the adjusted probability measure in π^* and the specification for $A(t,T)$ and $B(t,T)$

$$\frac{dA(t,T)}{dt} = -B(t,T)d^* - \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} B_i(t,T)B_j(t,T)\alpha_{ij} \tag{1.23}$$

and

$$\frac{dB(t,T)}{dt} = \iota' - B(t,T)C^* - \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} B_i(t,T)B_j(t,T)\beta_{ij} \tag{1.24}$$

Recall that we started with a factor process that is affine in the factor levels and also we have imposed an affine factor specification for the yields. These assumptions may seem arbitrary, but in fact Duffie and Kan (1996) show that an affine factor process implies an affine yield curve model and vice versa. For a proof the reader is referred to their work. Since the relation in (1.22) has to hold for all values of X_t and because the instantaneous interest (which appears on the right-hand-side) is affine in the factors, the factor drift and covariance (which appear on the left-hand-side) have to be affine in the factors as well, as was already imposed.

1.8 Exogenous no-arbitrage models

A sometimes unsatisfactory property of the endogenous models that have been described in the previous section is that these models do not exactly fit today's yield curve, which would allow for arbitrage opportunities. To deal with this undesirable property of endogenous models, Heath, Jarrow and Morton (1992, HJM) designed a class of models that exactly fits today's yield curve. Frachot and Lesne (1993) have introduced a general affine model specification within the HJM class. In essence, the entire current yield curve is the state variable, and every initial yield curve is consistent with a given parametrization of the model. In industry practice this is often handled by calibration of the current yield curve, where coefficients are allowed to vary in time to match the given initial yield curve. It is sufficient to adopt a time varying component in the drift part of the factor diffusion process. The endogenous models are just a special case of the exogenous models, where this component is restricted to be constant.

Following HJM we derive the term structure model in terms of forward rates. The implications for the yield curve follow after a simple transformation. With $f(t,T)$ we denote the forward rate for time T as seen from time t. The diffusion process under the risk adjusted probability measure is of the form

$$df(t,T) = m(t,T)\,dt + s(t,T)\,dW_t^* \tag{1.25}$$

where $m(t,T)$ is a scalar and represents the drift part, $s(t,T)$ is the $(1 \times K)$ vector of volatility factors, and W_t^* denotes a $(K \times 1)$ Brownian motion process. A well-known result from HJM is that the drift function is totally determined by the volatility specification

$$m(t,T) = s(t,T)\left[\int_t^T s(t,u)\,du\right]' \tag{1.26}$$

Substitution of (1.26) in (1.25) leads to the observation that the forward rate process is totally described by the volatility curve. We parameterize the volatility specification and also adopt a factor process that governs the dynamics of the forward rate curve. Assume that the volatility specifications depend on the underlying factor process and on time-to-maturity, hence

$$s(t,T) = s(X_t, T-t) \tag{1.27}$$

This leads to a parsimonious parametrization of the entire volatility curve since K factor take account for the dynamics of the entire volatility curve, where X_t denotes the $(K \times 1)$ underlying factor process, that takes account for the stochastic behavior of volatilities. Deterministic volatility is a special case when volatility does not depend on the factors.

An affine factor structure is incorporated for the forward rate curves, where K factors take account for the behavior of the entire forward rate curve

$$f(t,T) = a(t,T) + b(t,T)\,X_t \tag{1.28}$$

where $a(t,T)$ denotes a scalar, and $b(t,T)$ is a $(1 \times K)$ vector of parameters. The forward rates are related to the prices of discount bonds, given in equation (1.10), by

$$P(t,T) = \exp\left[-\int_t^T f(t,u)\,du\right] \tag{1.29}$$

which relates the coefficients of the discount bond representation and the forward rate representation in the following way

$$A(t,T) = -\int_t^T a(t,u)\,du \tag{1.30}$$

$$B(t,T) = -\int_t^T b(t,u)\,du \tag{1.31}$$

The factors again follow an affine diffusion process under the risk adjusted probability measure

$$dX_t = [C^* X_t + d^*(t)]\,dt + \nu(X_t)\,dW_t^* \tag{1.32}$$

Notice that, in comparison with the factor process in equation (1.6), the process is slightly different in that we have introduced a time varying scalar, $d^*(t)$, in the drift specification of the factors. The factor covariance specification is affine and the same as in equation (1.3). Applying Itô's lemma for the forward process in equation (1.28) and using the factor process in equation (1.32) yields the following specifications for both drift and volatility specification

$$m(t,T) = \frac{da(t,T)}{dt} + \frac{db(t,T)}{dt}X_t + b(t,T)[C^* X_t + d^*(t)] \tag{1.33}$$

and

$$s(t,T) = b(t,T)\,\nu(X_t) \tag{1.34}$$

Substitution of the volatility specification of equation (1.34) in (1.26) gives a second expression for the drift part.

$$m(t,T) = -b(t,T)\,V(X_t)\,B(t,T) = -\frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} b_i(t,T)\,B_j(t,T)(\alpha_{ij} + \beta_{ij}X_t) \tag{1.35}$$

Equating the drift specifications from equation (1.33) and (1.35) yields

$$\frac{da(t,T)}{dt} + \frac{db(t,T)}{dt}X_t + b(t,T)[C^* X_t + d^*(t)] = -\frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} b_i(t,T)\,B_j(t,T)(\alpha_{ij} + \beta_{ij}X_t) \tag{1.36}$$

or using the relations in (1.30) and (1.31)

$$\begin{aligned} -\left[\frac{dA(t,T)}{dt} - \frac{dA(t,T)}{dt}|_{T=t}\right] &= \left[\frac{dB(t,T)}{dt} - \frac{dB(t,T)}{dt}|_{\tau=0}\right]X_t + \\ &+ B(t,T)[C^* X_t + d^*(t)] + \\ &+ \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} B_i(t,T)\,B_j(t,T)(\alpha_{ij} + \beta_{ij}X_t) \end{aligned} \tag{1.37}$$

which results in the following Ricatti equations

$$\frac{dA(t,T)}{dt} = \frac{dA(t,T)}{dt}|_{T=t} - B(t,T)\,d^*(t) - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_{ij}B_i(t,T)\,B_j(t,T) \tag{1.38}$$

and

$$\frac{dB(t,T)}{dt} = \iota' - B(t,T)\,C^* - \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K}B_i(t,T)\,B_j(t,T)\,\beta_{ij} \tag{1.39}$$

Under the boundary conditions of equation (1.15) and (1.16) this results in almost the same Ricatti equations as in the endogenous model of the previous section. The only difference is the fact that here we have a time-varying drift, $d^*(t)$. The time-varying drift is specified in such a way that the initial term structure is modeled without error. Following Frachot and Lesne (1993), we impose that for each observed forward rate the factors are equal to the forward rates, such that the initial forward rate curve is modeled without error

$$f(t,t+\tau_i) = a(t,t+\tau_i) + \sum_{j=1}^{K} b_j(t,t+\tau_i)\,F_{jt}$$

with

$$\begin{aligned} a(t,t+\tau_i) &= 0 \quad i = 1,\ldots,K \\ b_i(t,t+\tau_i) &= 1 \quad i = 1,\ldots,K \\ b_j(t,t+\tau_i) &= 0 \quad i \neq j \end{aligned}$$

To accomplish this, the number of factors are actually equal to the number of observed forward rates, $K = N$. It follows immediately that, under this specification, the expected value of X_{it} is equal to the expected value of the observed forward rates, $f(t,t+\tau_i)$. Let

$$\psi(t) = E(X_t\,|X_0)$$

then for the forward rates it holds

$$\psi(t) = \begin{bmatrix} f(0,t+\tau_1) \\ \vdots \\ f(0,t+\tau_K) \end{bmatrix} + \begin{bmatrix} \int_0^t s(\psi(u),t+\tau_1-u)\left[\int_u^{t+\tau_1} s(\psi(u),v-u)\,dv\right]'du \\ \vdots \\ \int_0^t s(\psi(u),t+\tau_1-u)\left[\int_u^{t+\tau_1} s(\psi(u),v-u)\,dv\right]'du \end{bmatrix} \tag{1.40}$$

and from equation (1.32), it follows that

$$\psi(t) = \exp(-C^*t)\,X_0 + \int_0^t \exp[-C^*(t-s)]\,d^*(s)\,ds \tag{1.41}$$

Setting equations (1.40) and (1.41) equal to each other yields the time-varying drift part that exactly fits the initial yield curve.

1.9 Endogenous equilibrium models

Cox, Ingersoll and Ross (1985a, CIR) set up an intertemporal equilibrium model which is stated in continuous time and is applicable to derive asset prices in terms of the underlying variables describing the economy. The model endogenously determines the stochastic

process followed by the equilibrium price of any financial asset and shows that this process depends on the underlying real variables. Subsequently in Cox, Ingersoll and Ross (1985b) this framework is adopted to derive an equilibrium model for the term structure of interest rates. The framework as proposed by CIR is easily extended to derive other term structure models which belong to the affine class. Following Berardi and Esposito (1996) we show that in fact all models in the affine class are special cases of this general equilibrium model. Starting point is the model of CIR which we re-derive in order to set up the general framework.

Suppose there exist K state variables that describe the yield curve. By Z_{it} we denote the value of the i-th state variable at time t. We consider the case where the state variables belong to the affine class, which means that both the mean and the variance of the diffusion process are affine functions of the state variables

$$dZ_{it} = \xi_{0i}(\xi_{1i} - Z_{it})\,dt + \sqrt{\gamma_{0i} + \gamma_{1i}Z_{it}}dW_{it} \quad i = 1, \ldots, K \tag{1.42}$$

where ξ_{0i} is the mean reversion parameter, ξ_{1i} is the long term mean, γ_{0i} and γ_{1i} are the volatility parameters and W_{it} denotes Brownian motion. Well-known properties of the diffusion process are (i) if $\gamma_{1i} > 0$ then negative values for the state variables are precluded; (ii) if the state variable reaches zero, it can subsequently become positive; (iii) the absolute variance of the state variable increases when the state variable itself increases; (iv) there is a steady state distribution for the state variable. This specification encompasses as special cases the Gaussian model where $\gamma_{1i} = 0$ and the square root process where $\gamma_{0i} = 0$. We assume that the state variables are mutually uncorrelated.

In the economy that is set up, there exists a single production good which may be allocated to consumption or investment and all values are measured in terms of this good. The economy consists of three markets, a production market, a market for borrowing and lending and a derivatives market. The production possibilities consist of a set of N linear activities, which are proportional to the state variables. Let η denote the amount of the good invested in the production process, then the rate of return on the production process is of the following form

$$\frac{d\eta}{\eta} = \sum_{i=1}^{K} \mu_i Z_{it}\,dt + \sum_{i=1}^{K} \delta_i \sqrt{\gamma_{0i} + \gamma_{1i}Z_{it}}\,dW_{it} \tag{1.43}$$

where μ_i and δ_i are parameters that describe how the drift and diffusion of the return on the production process are related to the i-th state variable. This framework includes both uncertain production and random technological change, because the parameters of the diffusion process depend upon the underlying state variables which are random variables as well. There is a market for instantaneous borrowing and lending at interest rate r_t. The market clearing rate, as a function of the underlying variables, is determined as a part of the competitive equilibrium of the economy.

In the economy there are a number of identical agents which seek to maximize utility. In CIR (1985b) the preference structure is specialized to the case of time additive logarithmic

utility functions. Each individual seeks to maximize an objective function of the form

$$E\int_0^T e^{-\rho s}\ln c_s ds \tag{1.44}$$

where c_s denotes the consumption flow at time s. In this special case formulas simplify considerably in comparison with the general asset pricing formulas in CIR (1985a), since this specification reduces to the case where neither the interest rate r_t, nor the factor risk premiums depend on wealth. In equilibrium the wealth of individuals must be completely invested in the physical production process. The diffusion process which describes the dynamics of wealth, w_t, is given by[2]

$$\frac{dw_t}{w_t} = \frac{dq_t}{q_t} - \frac{c_t}{w_t}dt = \left[\sum_{i=1}^{K}\mu_i Z_{it} - \frac{c_t}{w_t}\right]dt + \sum_{i=1}^{K}\delta_i\sqrt{\gamma_{0i}+\gamma_{1i}Z_{it}}dW_{it} \tag{1.45}$$

The individuals choose optimal consumption and investment in the production process. Equilibrium in the economy determines the interest rate level, the contingent claim returns and the specification of the market price of risk. In equilibrium the net supply on riskless lending is zero, because we deal with identical individuals.

CIR (1985b) show that in equilibrium the real interest rate is given as the sum of K independent factors, X_{it}, $i = 1, \ldots, K$, which are linear transformations of the original factors:

$$r_t = \sum_{i=1}^{K} X_{it} \tag{1.46}$$

with

$$X_{it} = (-\delta_i^2\gamma_{0i}) + (\mu_i - \delta_i^2\gamma_{1i})Z_{it} \quad i = 1, \ldots, K \tag{1.47}$$

Because these new factors are a linear transformation of the state variables, Z_{it}, they also follow an affine diffusion process

$$dX_{it} = (C_i X_{it} + d_i)\,dt + \sqrt{\alpha_{ii} + \beta_{ii}X_{it}}dW_{it} \quad i = 1, \ldots, K \tag{1.48}$$

where C_i and d_i are scalars that determine the drift process of the i-th factor, and α_{ii} and β_{ii} are the volatility parameters. Applying the second theorem from CIR (1985a) for time additive logarithmic utility, implies that the equilibrium expected return on any contingent claim, G_t, is of the form

$$(\mu_G - r_t)G_t = \sum_{i=1}^{K}\frac{\partial G_t}{\partial X_{it}}\Phi_{it} \tag{1.49}$$

with market prices of risk given by

$$\Phi_{it} = cov\left(\frac{dW_t}{W_t}, dX_{it}\right) = \lambda_{0i} + \lambda_{1i}X_{it} \qquad i = 1, \ldots, K \tag{1.50}$$

[2]In case $\gamma_{1i} \leq 0$ this relationship may not lead to a sensible economy, since then state variables may become negative. This may result in negative interest rates, negative consumption and negative wealth. Hence, in order to arrive in a sensible description of the economy additional restrictions on the parameters are required.

where

$$\lambda_{0i} = \frac{\delta_i \alpha_{ii}}{\mu_i - \delta_i^2 \gamma_{1i}} \qquad i = 1, \ldots, K \tag{1.51}$$

and

$$\lambda_{1i} = \frac{\delta_i \beta_{ii}}{\mu_i - \delta_i^2 \gamma_{1i}} \qquad i = 1, \ldots, K \tag{1.52}$$

The risk-adjusted process for the factors is given by

$$dX_{it} = [d_i^* + C_i^* X_{it}]\, dt + \sqrt{\alpha_{ii} + \beta_{ii} X_{it}}\, dW_{it}^* \qquad i = 1, \ldots, K \tag{1.53}$$

where $d_i^* = (d_i - \lambda_{0i}\alpha_{ii})$, $C_i^* = (C_i - \lambda_{1i}\beta_{ii})$, and W_{it}^* denotes the risk adjusted Brownian motion. In the Gaussian case, $\beta_{ii} = 0$, the market price of risk is a constant as follows from equation (1.52) ($\lambda_{1i} = 0$) and in the square root case, $\alpha_{ii} = 0$, the market price of risk is linear in the factor, see equation (1.51) ($\lambda_{0i} = 0$). Because the factor processes in equation (1.48) are identical to the factor processes in the no-arbitrage case as in equation (1.6), imposing the exponential affine structure for discount bond prices as in equation (1.10), and taking the discount bond price as the derivative security, G_t, leads to the same Ricatti equations as in equations (1.23) and (1.24).

1.10 Solutions to the Ricatti equations

Aim of the previous sections has been to show how the parameters of an affine discount bond representation are related to the parameters of the associated factor process under the risk adjusted probability measure. This has resulted in Ricatti equations for the constants and factor loadings, given in equations (1.23) and (1.24), and in equations (1.38) and (1.39). In the first and second approach, we have adopted a general affine framework, where we allow the factors to be dependent. In the third approach, we have implicitly assumed that the factors are uncorrelated. In the following, we present a solution to the Ricatti equations for the case where the factors are assumed to be independent.

Assume that the constant term in a multifactor model, $A(t,T)$, equals the sum of the constant terms in single factor models

$$A(t,T) = \sum_{k=1}^{K} A_k(t,T) \tag{1.54}$$

where $A_k(t,T)$ is a constant related to the k-th factor. The price for a single factor discount bond, that is affine in the factor X_{kt}, then follows as

$$P_k(t,T) = \exp\left[A_k(t,T) + B_k(t,T) X_{kt}\right] \quad k = 1, \ldots, K \tag{1.55}$$

where $B_k(t,T)$ is the k-th element of the vector of factor loadings, $B(t,T)$. This implies that the price of a multifactor affine discount bond is given by the product of the K single factor discount bonds in equation (1.55)

$$P(t,T) = \prod_{k=1}^{K} P_k(t,T) \tag{1.56}$$

The independent factors specification simplifies the solution of the Ricatti equations considerably, since instead of solving the equations for $A(t,T)$ and the vector $B(t,T)$, they can be expressed in terms of $A_k(t,T)$ and $B_k(t,T)$ and hence solved in the single factor case. The multifactor expressions follow simply from equation (1.56). The Ricatti equations for the single factor bond prices read

$$\frac{dA_k(\tau)}{d\tau} = B_k(\tau)\, d_k^* + B_k^2(\tau)\, \alpha_{kk} \tag{1.57}$$

and

$$\frac{dB_k(\tau)}{d\tau} = -1 + B_k(\tau)\, C_k^* + B_k^2(\tau)\, \beta_{kk} \tag{1.58}$$

Solving first Ricatti equation (1.58) and then equation (1.57) yields a closed form solution for price $P_k(t,T)$ in equation (1.55). An expression for the price of the discount bond that depends on multiple factors follows easily as the product of the solutions for the single factor discount bond expressions, as given in equation (1.56). Solutions for the Ricatti equations in (1.57) and (1.58) are given by

$$A_k(\tau) = \ln\left\{\left[\frac{2\gamma_k e^{\left[\frac{1}{2}(\gamma_k - C_k^*)\tau\right]}}{(\gamma_k - C_k^*)(e^{\gamma_k\tau} - 1) + 2\gamma_k}\right]^{2\left[\frac{d_k^*}{\beta_{kk}} - C_k^*\frac{(\alpha_{kk}+\beta_{kk})}{\beta_{kk}}\right]} e^{\left[-\frac{\alpha_{kk}}{\beta_{kk}}(B_k(\tau)-\tau)\right]}\right\} \tag{1.59}$$

and

$$B_k(\tau) = -\frac{2(e^{\gamma_k\tau} - 1)}{(\gamma_k - C_k^*)(e^{\gamma_k\tau} - 1) + 2\gamma_k} \tag{1.60}$$

with

$$\gamma_k = \sqrt{(C_k^*)^2 + 2\beta_{kk}} \tag{1.61}$$

In the term structure literature two basic single factor models are distinguished, the CIR model and the Vasicek model. Because multifactor yield curve models are simply the addition of single factor yield curve specifications, the CIR and Vasicek models basically serve as building blocks of more elaborate term structure models. The CIR model is a special case of the general affine term structure model, specified by $\alpha_{kk} = 0$, and $\lambda_{0k} = 0$. The factor process under the risk adjusted probability measure reads

$$dX_{kt} = [d_k + C_k^* X_{kt}]\, dt + \sqrt{\beta_{kk} X_{kt}}\, dW_{kt}^* \tag{1.62}$$

The volatility part is proportional to the factor level, which takes account for stochastic volatility. Because $\lambda_{0k} = 0$, it holds that $d_k = d_k^*$. The constant in the drift part is equal to the constant under the actual probability measure. The expression for $A_k(\tau)$ simplifies to

$$A_k(\tau) = \ln\left\{\left[\frac{2\gamma_k e^{\left[\frac{1}{2}(\gamma_k - C_k^*)\tau\right]}}{(\gamma_k - C_k^*)(e^{\gamma_k\tau} - 1) + 2\gamma_k}\right]^{2\left[\frac{d_k^*}{\beta_{kk}} - C_k^*\right]}\right\} \tag{1.63}$$

whereas the expressions for $B_k(\tau)$ and γ_k are the same as in the general case. The Vasicek model is another special case of the general affine model, specified by $\beta_{kk} = 0$, and $\lambda_{1k} = 0$. The factor process reads

$$dX_{kt} = [d_k^* + C_k X_{kt}]\, dt + \sqrt{\alpha_{kk}} dW_{kt}^* \tag{1.64}$$

The factor process is Gaussian and volatility is constant, since it no longer depends upon the factor level. Because $\lambda_{1k} = 0$, it holds that $C_k = C_k^*$. The slope term in the drift part now is identical to its value under the actual probability measure. The Ricatti equation in (1.58) simplifies considerably and has the following solution

$$B_k(\tau) = -\frac{(1 - \exp[C_k \tau])}{C_k} \tag{1.65}$$

Substitution of (1.65) in equation (1.57) yields the following expression for $A_k(\tau)$ in the Vasicek case

$$A_k(\tau) = \ln[(\tau - B_k(\tau))\left(\frac{d_k^*}{C_k} + \frac{1}{2}\frac{\alpha_{kk}}{C_k^2}\right) + \alpha_{kk}\frac{\left(1 - e^{C_k \tau}\right)^2}{4C_k^3}] \tag{1.66}$$

Multifactor models differ in the number of underlying factors and in the type of stochastic processes for the dynamics of these factors. Often the factors as incorporated in the affine class are unobservable. Berardi and Esposito (1997) show that when term structure models belong to the time homogeneous class of affine term structure models, they are easily derived by means of simple reshuffling of the base model, in which the price of a discount bond or the associated yield curve is modeled as an affine function of unobserved factors. They show that basically two methodologies are applicable to express the unobserved factors in terms of observables, as proposed by the literature. The first methodology is defined as 'yield-based inversion' and consists of expressing the state variables as linear functions of (a subset) of yields. The second methodology is defined as 'shape-based inversion'. This second methodology enables us to express the factors as more general functions of observables.

The 'yield-based inversion' makes use of the affine relation between yields and factors. In an affine term structure model the relation between yields and factors is given by

$$-(T-t)\,Y(t,T) = \left[A(T-t) + \sum_{k=1}^{K} B_k(T-t)\, X_{kt}\right] \tag{1.67}$$

Provided the assumption that the number of yields (N) is greater than the number of factors (K), each factor can be written as a linear combination of yields. To accomplish this, just pick K points at the yield curve, $T_1, \ldots, T_K$, and express them in terms of the affine model of equation (1.67)

$$\begin{array}{rcccccccc}
-(T_1-t)\,Y(T_1-t) - A(T_1-t) & = & B_1(T_1-t)\,X_{1t} & + & \cdots & + & B_K(T_1-t)\,X_{Kt} \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
-(T_K-t)\,Y(T_K-t) - A(T_K-t) & = & B_K(T_K-t)\,X_{1t} & + & \cdots & + & B_K(T_K-t)\,X_{Kt}
\end{array}$$

Solving the system of equations yields an expression for the unobservable factors in terms of the observed yields on the curve. Hence, other yields may be rewritten in terms of these yields, that replace the factors. Duffie and Kan (1996), Brennan and Schwartz (1979), Chen and Scott (1993) and El Karoui and Lacoste (1992) are examples where this methodology has been applied. Since the spread between two yields is also a linear combination of yields, models as proposed by Schaefer and Schwartz (1984) and Brown and Schaefer (1994) also belong to the 'yield-based inversion' class.

Models that belong to the 'shape-based inversion' are concerned with variables that measure the curvature of the term structure. Examples are forward rates and volatility of yields. Forward rates are defined by

$$f(t,T) = -\frac{\partial \ln[P(t,T)]}{\partial T} \tag{1.68}$$

which gives the following expression in the case of affine term structure models

$$f(t,T) = r_t - \sum_{k=1}^{K} B_k(T-t)(d_k^* + C_k^* X_{kt}) - \frac{1}{2}\sum_{k=1}^{K} B_k(T-t)^2(\alpha_{kk} + \beta_{kk} X_{kt}) \tag{1.69}$$

Forward rates are affine in the factors and depend on both the drift parameters and the volatility parameters of the factor process. Similar to the 'yield-based inversion' methodology, we are able to express the factors as linear combinations of a particular set of forward rates. The remaining forward rates can therefore be expressed as a function of these basis forward rates. Application of this technique leads to models as proposed by Stambaugh (1988), El Karoui and Lacoste (1992), Frachot and Lesne (1993) and an affine specification of the HJM (1992) class.

Another class of 'shape-based inversion' arises when we incorporate volatility of yields as factors. Recall that the short rate is the sum of the factors

$$r_t = \sum_{k=1}^{K} X_{kt} \tag{1.70}$$

With independent factors this gives that the variance of the short rate is equal to the factor variances

$$\sigma_t^2 = \sum_{k=1}^{K} (\alpha_{kk} + \beta_{kk} X_{kt}) \tag{1.71}$$

where σ_t^2 denotes the variances of the short rate at time t. In order to set up an inversion table we need as many equations as factors. A second linear relation between factors and economic variables comes from the short rate. Other linear relations between factors and economic variables are available by the inclusion of variances of longer term yields

$$\mathrm{Var}(Y(t,T)) = \left(\frac{1}{T-t}\right)^2 \sum_{k=1}^{K} B_k(T-t)^2(\alpha_{kk} + \beta_{kk} X_{kt}) \tag{1.72}$$

Inclusion of these relations are interesting because they focus on different moments in the data than in the 'yield-based inversion' methodology. For bond option pricing the inclusion

of volatility terms is particular important. For longer term derivatives it seems appropriate to include volatility of longer term yields, instead of using the volatility of the short rate, which is common practice in time series approaches of interest rates. Examples of this 'shape-based inversion' methodology include Longstaff and Schwartz (1992a), who adopt the short rate and the volatility of the short rates as the two factors driving the yield curve.

1.11 Research objectives

In classical approaches to model the term structure of interest rates, the focus is on estimating the underlying factor specifications. Usually, we require some approximation for the factors since they are not directly observable. An example is the instantaneous spot rate which is often approximated by some short rate like the one month or the three month T-bill rate. The fact that the relevant factors are often not directly observable leads to an approximation bias, which may influence the results. Moreover, since the process of the underlying factors is under the actual probability measure we require a market price of risk to infer term structure implications. The properties of the market price of risk are however not very well understood and usually assumptions are made with respect to its functional representation. Finally, a model for the underlying factors is estimated and not for the term structure directly. So, even though the model may properly represent the underlying factors, this does not mean that the term structure is adequately described as well. Therefore, focusing on the underlying factors can have some serious drawbacks.

The drawbacks of the aforementioned approach serves as a motivation for this dissertation. We propose to employ a panel data analysis of the term structure of interest rates. In our view this approach is natural for a number of reasons. First, it seems logical to use yield curve data when we estimate models of the term structure of interest rates, since these models describe the yield curve. This prevents us from approximation bias in modeling the underlying factors. Second, because there exist a lot of coupon bearing bonds that are informative on the term structure of interest rates, there is an efficiency gain in comparison with the factor approach. Third, no approximations have to be made about the market price of risk, since for yield curve data only the risk adjusted process is important. In general, for the pricing of any derivative security only the risk adjusted processes are relevant.

In a panel data analysis all available data is incorporated simultaneously to estimate the model. Analysis under the actual probability measure deal with a panel of time series, that focuses on the dynamics of all yields simultaneously. Analysis under the risk adjusted probability measure considers the cross sectional information in the yield curve. All cross sections of yield curves are pooled to estimate the model.

To simplify the use of panel data techniques on term structure of interest rates, it is desirable to have a regular panel of yields as data. Yields are not directly observable, but are implied by bond prices. The analysis of term structure models on coupon bearing

bond prices directly leads to non-linearities, since bond prices are non-linear in yields. In chapter two we discuss alternative methods to construct a panel of discount yields from bond prices. There we will compare the properties of different functional specifications and different econometric techniques. The constructed panel of yields will be used in all the subsequent analyses.

In chapter three we start to investigate the time series information that is available in the data. We consider panel data implications in a model free framework, i.e. we do not adopt a specific model structure to represent the yield curve. Panel data methods are applied to the well-known problem of testing the expectations model. The focus is on the risk premia in the term structure of interest rates. In particular the panel data framework allows us to explicitly model the dynamics of the risk premium. We deal with both the cross sectional properties of risk premia, i.e. how are the risk premia for bonds with different maturities related to each other, and with the time series dynamics of the risk premium, i.e. how do risk premia vary over time. This second feature is established by assuming that the risk premium is a common factor in yields of all maturities. Term premia are interesting, because when they can vary, monetary policy can affect real long term rates through other channels than just by the short term rate. The term premium is also closely related to the market price of risk, which is an important ingredient for derivatives pricing.

In chapter four we investigate the cross sectional information in the data. We consider the class of affine term structure models. Yields are expressed in terms of parameters under the risk adjusted probability measure directly, which leads to testable cross sectional restrictions. We build a general econometric framework that encompasses both exogenous and endogenous term structure models and also single and multifactor models. Since we employ panel data techniques two features are important. First, the analysis is simplified considerably when we use a regular panel of yields as constructed in chapter two. Second, we need a tractable model to deal with the factors that appear as explanatory variables in the model. We do not want to incorporate time series information of the underlying factors because this requires an assumption on the market price of risk. The alternative is to treat the factors as unobserved parameters that are estimated along with the structural parameters of the model. This means that in a K factor, with a time series of T observations, KT parameters are added to model. We use a model that is affine in the underlying factors, since then standard panel data techniques exist to deal with the factors. Another desirable feature of affine term structure models is that they lead to closed form solution of the yield curve.

A second important feature of chapter four concerns the motivation for the error term that we add to the affine term structure model. We motivate at length the properties of its covariance matrix and also we include autocorrelation for the error terms. This leads to a parsimonious econometric specification that allows us to estimate a general class of models and to carry out all kinds of diagnostic tests.

In chapter five, we investigate the properties of the Longstaff-Schwartz term structure

model. The model belongs to the affine class and it adopts the instantaneous spot rate and the volatility of the instantaneous spot rate as underlying factors. From the literature it is known that the model specification of the underlying factors have been successful representations of the time series dynamics of these factors. Furthermore, the analysis in chapter four indicates that a two factor Gaussian term structure model leads to a good representation of yield curve. Hence, such a model is a good starting point for a cross sectional analysis. We compare term structure implications from factor process estimates, with implications from panel data estimates of the yield curve directly, because it is interesting to see whether both approaches lead to similar results or where they would differ. We also carry out an out-of-sample test, since we are interested whether difference in both approaches are also significant for the pricing of other derivative securities.

In chapter six, we provide an overview of the main findings of this dissertation and we give directions for further research.

Chapter 2

Curve Fitting Methods

2.1 Introduction

The term structure of interest rates can be represented by different economic variables, like bond prices, yields and forward rates. The prices of coupon bearing bonds are directly available, since they are actually quoted in the bond market. Yields and forward rates are derived from the bond prices. In the literature term structure models are often stated and derived in terms of discount bond and discount yields. A coupon bearing bond is considered to be a basket of discount bond instruments with different cashflows, i.e. the coupons, and with different times to maturity, i.e. the coupon payment dates. The price of a coupon bearing bond hence follows as the price of a number of discount bonds. The analytical treatment for a discount bond is easier than for a coupon bearing bond. Moreover, the relations between bond price, yield and forward rate is straightforward in the case of discount bonds.

The term structure of interest rates, or the yield curve, describes the relation between the implied return on a discount bond and its associated time to maturity. In practice, the yield curve is not directly observable. Only for short maturities discount yields are available. Information on the discount yield curve is however implicit present in bond data. In absence of arbitrage opportunities the price of a bond is the present value of all future coupon payments and of the principal at maturity of the bond. The coupon payments and the principal would be discounted at the current discount yield curve.

Since it will rarely be the case that exactly all time-to-maturities on the yield curve are available in bond data, a parametric function has often been specified in practice. Notice, that we observe only a finite number of bonds with a finite number of maturities, whereas the yield curve is a continuous curve. By specifying a function for the discount rates, we are able to fit the function at a discrete number of points that are observed and interpolate for all the points in between.

The motivation for doing this is threefold. First of all, we construct in a way discount yield data, which are required for empirical research of term structure models, like Vasicek (1977), Cox Ingersoll and Ross (1985) and the no-arbitrage models as described in Heath,

Jarrow and Morton (1992). These models are usually stated in terms of discount yields. Testing these models on bond data directly leads to non-linear relationships, since only for discount bonds the transformation from yields to prices is easy, where it is not for coupon bearing bonds.

One problem with using the pricing relation as stated above, is that it assumes a world of certain cash, and frictionless and complete markets. In such a world absence of arbitrage indeed is equivalent to a unique pricing rule. Bliss (1994) stresses that we need only look at the quotations of treasury strips to observe apparent violation of the law of one price, implying that the frictionless markets assumption is not supported by facts. Other studies have used linear programming methods to detect arbitrage opportunities, implying that a single discount rate does not exist, see for example Cornell and Shapiro (1989).

A second motivation for constructing yield curve functions from bond data, is that the differences between the observed prices and modeled prices quantify the impact of all the remaining effects, that are not included in the pricing relation. Since the relations explaining these residual effects are difficult and not entirely clear, no attempt is made to take account for them explicitly. The models in this paper only incorporate the pricing relation stated before, which means that all the remaining effects will be in the residuals. The term structure literature has sought to explain these phenomena in terms of frictions such as taxes, sale constraints, liquidity premia, transaction costs and non-synchronous trading. The residuals provides us with a set of data that contain information on the presence of these effects. One simple way to gain insight into the apparent relations is by running regressions of the residuals on variables which represent these frictions. The exercise of fitting a pricing relation to the data also provides us with a way to test how well such a pricing relation works in practice.

A third motivation lies in the fact that together with the parameter estimates for the curve fitting models, we obtain a covariance matrix for the parameter estimates. The functional specification is evaluated in the point estimates, which leads to the discount yield curve for a specific date. The covariance matrix of the parameter estimates allow us to quantify the uncertainty that is present in this constructed yield curve. In the empirical application confidence intervals represent this uncertainty.

In the next section we describe the econometric framework to construct a term structure of discount yields from bond data. We focus on two methods that have been applied in the literature, the cubic spline method of McCulloch (1971, 1975) and the Nelson and Siegel (1987) exponential approach. Also an extended Nelson-Siegel method is incorporated as proposed by Bliss (1994). It turns out that the estimation criterion has an important impact on the properties of the implied yields. Therefore, we investigate different estimation criterions for these models and compare the estimation results with each other. In a second stage we perform a residual analysis, to detect any missing explanatory variables and to come up with a measure of accurateness of the methods employed.

Of course, many other approximating method have been pursued in the literature, like polynomials (Chambers, Carleton and Cooper (1984)), step functions (Ronn (1987))

and piecewise linear functions (Fama and Bliss (1987)).[1] Also many different estimation procedure have been employed. Typically, the linear programming methods (Ronn (1987)) and the iterative extraction method (Fama and Bliss (1987)) focus on no-arbitrage or fitting the bond prices without error. Econometric methods like weighted least squares and maximum likelihood try to fit the data as good as possible for all the bond prices and, at the same time, preserve smoothness of approximating functions.

Fama and Bliss (1987) construct yield curve data for maturities of 1 year to 5 years, which are stored on the CRSP tapes. They use a number of filtering methods to select particular bonds to construct the yield curves from. It is unclear how to interpret these yield curves, since they only match with a particular number of bonds. We construct yield curves as well, but we do not leave out any of the data.[2] To provide the reader with a measurement for the quality of the specific time series, we also give confidence intervals for all yields. We regress the associated residuals on a number of factors to see what bonds are represented well by the yield curves and for which bond prices other effects are important.

Bliss (1994) develops and tests procedures for evaluating term structure estimation methods. The purpose of his paper is to establish a methodology to be used in the future to test whether more complex pricing relations fully capture the systematic behavior of bond prices. Essentially, he stresses that models should be evaluated at their out-of-sample performance. He does however not elaborate on the model that would be preferred.

An important issue is whether it is justified to use yields that are constructed by the application of some statistical model as input for other term structure models. All term structure models satisfy the law of one price that underlies the estimation technique we employ in this paper. However, because most term structure models impose more structure on the shape of the yield curve, they are more restrictive than the Nelson-Siegel model and the cubic spline method. This justifies the use of a two-step procedure to do term structure research. In the first step, discount yields are constructed from coupon-bearing bonds and in the second step these constructed yields are used to test the more restrictive term structure models. A model that is more restrictive can never fit the raw data better, so one might as well use the constructed discount yields.

2.2 Econometric framework

The current price of a bond is related to the future coupon payments and the payment of the principal when the bond matures. The discount function is an instrument that relates future cashflows to the current value of the bond. Since both bond prices and cashflow values are given, the discount function is implied by these two. At each time t we observe

[1] Bliss (1994) gives a good overview of the different approximating functions and of the different methods used to estimate these functions.

[2] We already restricted our data set to contain bonds, bills and notes. So in fact a lot of interest rate sensitive products have not been included. We are interested in a term structure for interest rate securities that have no option features and do not include all sorts of tax-features.

a set of coupon bearing bonds, which differ with respect to the level of the coupon, the coupon payment dates and the time-to-maturity. Also, the number of bonds that are available at each time are different. We focus on each cross section separately and extract a discount function from the available bond data at that time. For the econometric model four ingredients are required: an economic relationship, the inclusion of an error term, an approximate functional specification for the discount factor and some fitting criterion.

The pricing relation of a bond in a complete and frictionless market is stated as

$$P_i(t) + A_i(t) = c_i \sum_{j=1}^{N_i(t)} \delta(t, \tau_{ij}) + 100\delta(t, \tau_{i,N_i(t)}) \qquad i = 1, \ldots, M(t) \quad t = 1, \ldots, T \quad (2.1)$$

At time t we have $M(t)$ bonds at our disposal, where $P_i(t)$ is the price at time t, $A_i(t)$ is the accrued interest and c_i is the periodical coupon payment for bond i. The coupon payment dates of different bonds usually do not coincide and therefore we adopt a bond specific stream of coupon payment dates. At time t bond i has $N_i(t)$ remaining coupon payment dates, and the coupons are paid at times $t + \tau_{ij} \quad j = 1, \ldots, N_i(t)$. At time $t + \tau_{i,N_i(t)}$ the bond matures and the principal of 100 is retained. Cashflows at time $t + \tau$ are discounted with the discount functions $\delta(t, \tau)$. The discount function gives the present value at time t of a cashflow worth one dollar at time $t + \tau$. The relation in equation (2.1) holds for each bond at each time. In the estimation routine we consider cross sections of bonds, i.e. we use all the bonds that are available at a certain date. The relation between the discount function and the yield curve is given by

$$\delta(t, \tau) = \exp[-\tau Y(t, \tau)] \quad (2.2)$$

which implies a yield curve at time t of the form

$$Y(t, \tau) = -\frac{1}{\tau} \ln[\delta(t, \tau)] \quad (2.3)$$

Typically, the discount functions are unknown, whereas the remaining variables are observed. A popular method to infer the discount function from the bond prices is by specifying some parametric form for it.

In practice the relation as stated in equation (2.1) does not hold, since markets are not frictionless. Furthermore, we do not observe a single price, but in fact prices are quoted within a spread, the difference between an askprice, $P_i^A(t)$ and a bidprice, $P_i^B(t)$. Dermody and Prisman (1988) argue that, because of this spread, multiple discount functions exist, satisfying

$$P_i^B(t) \leq c_i \sum_{j=1}^{N_i(t)} \delta(t, \tau_{ij}) + 100\delta(t, \tau_{i,N_i(t)}) - A_i(t) \leq P_i^A(t) \quad (2.4)$$

which means that all prices in between the bid- and the askprice are admissible. Usually the midprice, $P_i^M(t)$, which is the average of the bidprice and the askprice is used in empirical work, see for example McCulloch (1971) and Nelson-Siegel (1985). We incorporate an error term which takes account of all the effects that are not in the pricing relation. There are

two possible ways to model and interpret an error term within this framework. In the first case, the error term is the difference between the observed (mid-)price and the modeled price. This results in an econometric model of the form

$$P_i^M(t) + A_i(t) = c_i \sum_{j=1}^{N_i(t)} \delta(t, \tau_{ij}) + 100\delta(t, \tau_{i,N_i(t)}) + \varepsilon_i(t) \qquad i = 1, \ldots, M(t) \quad t = 1, \ldots, T \tag{2.5}$$

where $\varepsilon_i(t)$ denotes the error term that captures all the remaining effects. In Bliss (1994) a second approach is pursued to incorporate an error term. Instead of using the midprice as the single dependent variable, any modeled price between the bid- and askprice is considered as a perfect fit of the observed price. The error term is therefore defined to be zero for any modeled price in between the bid- and askprice. Outside this interval the error term is defined as the difference between the bid/ask price and the fitted price. Consider the regression equation

$$P_i^M(t) = \hat{P}_i(t) + \eta_i(t) \tag{2.6}$$

where $\hat{P}_i(t)$ is the modeled price, given by

$$\hat{P}_i(t) = c_i \sum_{j=1}^{N_i(t)} \delta(t, \tau_{i,j}) + 100\delta(t, \tau_{i,N_i(t)}) - A_i(t) \qquad i = 1, \ldots, M(t) \quad t = 1, \ldots, T \tag{2.7}$$

and the error term, $\varepsilon_i(t)$, is defined in terms of the help variable $\eta_i(t)$ as

$$\varepsilon_i(t) = \begin{cases} \hat{P}_i(t) - P_i^A(t) & \eta_i(t) < P_i^M(t) - P_i^A(t) \\ \hat{P}_i(t) - P_i^B(t) & \eta_i(t) > P_i^M(t) - P_i^B(t) \\ 0 & \text{otherwise} \end{cases} \tag{2.8}$$

The next thing to specify is the exact functional form that is used to approximate the discount function. Examples of the approximating functions include polynomials, cubic splines, piecewise linear functions and exponential forms. Vasicek and Fong (1982) state that the objective in empirical estimation is to find a discount function that (a) fits the data sufficiently well, and (b) is a sufficiently smooth function. This second requirement is less often stated. It ensures that yield curves are smooth functions, which means that the difference in discount rates at different maturities, say, τ and $\tau + d\tau$, are small when $d\tau$ is small. Furthermore, it ensures that the yield curve is differentiable in each point, which is a desirable property to determine instantaneous forward rates from the yield curve.

Especially the cubic splines and the exponential forms are in accordance with the requirements as stated by Vasicek and Fong (1982). In the empirical section we will deal with the cubic spline method as proposed by McCulloch (1971, 1975) and with the exponential specification, developed by Nelson and Siegel (1987). These functional specifications are described at length in the next section.

Last thing to do is to decide upon the method for estimating the parameters of the approximate functions. The cubic spline model is attractive since it is linear in the parameters and hence some weighted least squares method seems appropriate. The Nelson-Siegel

function is non-linear in the parameters, but still a non-linear (weighted) least squares estimation procedure is applicable. In the term structure literature two methods have been proposed to weight the errors. First of all, the spread, the difference between the bid- and askprice, may reflect the dispersion in bond prices, hence the normalized reciprocal of the spreads are plausible weights.[3] The spread of bond i at time t is defined as

$$v_i(t) = P_i^A(t) - P_i^B(t) \tag{2.9}$$

and so at time t, the weight of bond i is given by

$$w_i(t) = [1/v_i(t)] / \left[\sum_{j=1}^{M(t)} 1/v_j(t)\right] \tag{2.10}$$

Secondly, the duration of the bond reflects uncertainty with respect to its price. Long term bonds consist of future coupon payments, which are all discounted with error. The error of the bond price cumulates with its duration, therefore the normalized reciprocal of the duration of the bonds, are used as weights, given by

$$w_i(t) = [1/D_i(t)] / \left[\sum_{j=1}^{M(t)} 1/D_j(t)\right] \tag{2.11}$$

where $D_i(t)$ denotes the duration of bond i at time t.[4] Following the argument of Bliss (1994), the squared error terms as defined in equation (2.8) is minimized, possibly weighted with either the inverse of the spreads, or with the inverse of the durations. This results in an objective function at time t of the form

$$\sum_{i=1}^{M(t)} w_i(t)\,\varepsilon_i^2(t) \tag{2.12}$$

which is minimized in order to find the discount function that best describes the bond data at time t.

2.3 Approximate function specification

Without any assumptions the discount rates are not identified from the bond prices. Since cashflow payments occur in general at different times for each bond, there are more discount

[3]The weights are actually normalized, such that they are all positive and add up to one. This scaling is helpful in avoiding all sorts of numerical problems that can arise when the weights are of a different order than the prices and the coupons.

[4]Similar to Bliss (1994) we found that duration weighting is much more successful than spread weighting or equal weighting. In the latter two cases we found extreme outliers and had to filter the data to come up with reasonable implied term structures. Using maturity instead of duration almost gives the same results, but duration better distinguishes the early coupon payments, and hence better captures interest rate risk.

rates than bond prices. A common approach is to adopt some functional form for the discount function, that on the one hand is flexible, so it is able to fit the observed bond prices accurately, and on the other hand is parsimonious in the number of parameters to be estimated. We describe the cubic spline function as proposed by McCulloch (1971, 1975), the Nelson-Siegel (1987) exponential functions and the extended Nelson-Siegel version as proposed by Bliss (1994).

A cubic spline function gives a functional form of the discount function like

$$\delta(t,\tau) = 1 + \alpha_1(t)\tau + \alpha_2(t)\tau^2 + \sum_{k=1}^{K(t)} \alpha_{2+k}(t)\left[\tau - b_k(t)\right]_+^3 \tag{2.13}$$

where $\alpha_l(t)$, $l = 1, \ldots, K(t)+2$ are unknown parameters, $b_k(t)$, $k = 1, \ldots, K(t)$ are breakpoint which are set at appropriate values and $[\cdot]_+ = \max[\cdot, 0]$. Notice that $\delta(t,0) = 1$, since the present value of present money is unity. One way to choose the number of breakpoints, $K(t)$, is by doing a grid search over the range $[0, M(t)]$. That is, the number of breakpoints is between zero and the number of observations at a certain time t. Pick the value for $K(t)$ that produces the lowest value for $\hat{\sigma}_t^2$, with

$$\hat{\sigma}_t^2 = \frac{1}{M(t) - K(t) - 2} \sum_{i=1}^{M(t)} \left\{ w_i(t) \left[P_i^M(t) - \hat{P}_i(t) \right]^2 \right\} \tag{2.14}$$

where $\hat{\sigma}_t^2$ is the unbiased estimator of σ_t^2, the variance of the error term. For $w_i(t)$ we choose one of the specifications from equations (2.10) and (2.11). As $K(t)$ increases the residuals usually decrease, but then so do the degrees of freedom. The result is that $\hat{\sigma}_t^2$ declines sharply as $K(t)$ increases from 2 to 3 or 4, but thereafter fluctuates irregularly with a small amplitude, see McCulloch (1971). A second approach would be to make $K(t)$ a fixed function of the number of bonds, $M(t)$. For $K(t)$ to increase when $M(t)$ increases, it should be an increasing function. In order to make the number of observations in the domain of each quadratic segment increase with the total number of observations, the ratio $M(t)/K(t)$ should also increase with $M(t)$. A function that fulfills these properties is

$$K(t) = \left\lfloor \sqrt{M(t)} \right\rfloor \tag{2.15}$$

that is, the number of breakpoints should be chosen as the integer part of the square root of the number of bonds. The breakpoints divide the spectrum of maturities in equal parts, such that the function is flexible enough to take account for all parts of the yield curve. In practice, this formula gives approximately the same results as the grid approach, see McCulloch (1971). Notice that the choice for $K(t)$ is an important one, since setting it too low does not allow the function to be flexible enough, which will result in a worse fit. Setting $K(t)$ too high, would allow for the possibility of overfitting, or noise-fitting.

Substitution of equation (2.13) in equation (2.7) gives

$$\hat{P}_i(t) = [N_i(t)c_i + 100] + \alpha_1(t)\left[c_i \sum_{j=1}^{N_i(t)} \tau_j + 100\tau_{Ni(t)}\right] + \alpha_2(t)\left[c_i \sum_{j=1}^{N_i(t)} \tau_j^2 + 100\tau_{Ni(t)}^2\right] +$$

$$\sum_{k=1}^{K_t} \alpha_{2+k}(t) \left[c_i \sum_{j=1}^{N_i(t)} [\tau_j - b_k(t)]_+^3 + 100 [\tau_{Ni} - b_k(t)]_+^3 \right] - A_i(t) + \varepsilon_i(t) \qquad (2.16)$$

Notice that the model is still linear in the parameters, so (weighted) linear least squares techniques are applicable. In the empirical application we minimize the squared error term as given in equation (2.12), where we fix the number of breakpoints as given by equation (2.15). Typically, we will be estimating models with 8 to 15 breakpoints and hence 10 to 17 unknown parameters, since usually about 75 to 225 bond prices are available at each time t.

The implied instantaneous interest rate, which is the rate when time to maturity approaches zero, is given by

$$Y(t,0) = -\lim_{\tau \to 0} \frac{\ln[\delta(t,\tau)]}{\tau} = \alpha_1(t) \qquad (2.17)$$

and the infinite yield is given by

$$Y(t,\infty) = -\lim_{\tau \to \infty} \frac{\ln[\delta(t,\tau)]}{\tau} = \infty \qquad (2.18)$$

Hence the cubic spline yield curves diverge as time-to-maturity tends to become large. The instantaneous spot rate converges however. If the estimated values for $\alpha_1(t)$ for $t = 1, \ldots, T$ are estimated precisely, then it would provide us with an instantaneous spot rate that is implied by and consistent with the yield curve. Time series approaches to test term structure models are always stated in terms of the instantaneous spot rate (see Chan, Karolyi, Longstaff and Sanders (1992)). Since the instantaneous spot rate is not directly observable, often the 1 month observed spot rate is used as a proxy. The one month rate is however exposed to a lot of transitory noise, which might contaminate the results.

Nelson and Siegel (1987) have proposed a class of models, which is motivated by the expectations hypothesis and which offers a parsimonious description of the different shapes that have traditionally been associated with yield curves. Yield curves are typically monotonic increasing, flat, inverse, humped and occasionally S-shaped. Instead of specifying the discount function, Nelson and Siegel (1987) model the yield curve directly as

$$Y(t,\tau) = \beta_0(t) + [\beta_1(t) + \beta_2(t)] \left[\frac{1 - \exp[-\tau/\gamma(t)]}{\tau/\gamma(t)} \right] - \beta_2(t) \exp[-\tau/\gamma(t)] \qquad (2.19)$$

where $\beta_0(t)$, $\beta_1(t)$, $\beta_2(t)$ and $\gamma(t)$ are the unknown parameters to be estimated with the data available at time t. Notice that this is a more parsimonious setting than the cubic spline case, but it is still flexible enough to model the different shapes of a term structure.

To illustrate this they show that equation (2.19) consists of three components. The first component is equal to 1 and its contribution to the yield curve is measured by $\beta_0(t)$. This is the long term component since it has the same influence on each part of the yield curve. The second term is $[1 - \exp(-\tau/\gamma(t))] / [\tau/\gamma(t)]$ which is small for low maturities,

increases until some point and damps out when maturity is large, hence this is the term that deals with the mid part of the curve. The contribution of this medium term is measured by $\beta_1(t)+\beta_2(t)$. Lastly, the term $\exp[-\tau/\gamma(t)]$ is highest for low maturities and decreases exponentially fast as maturity rises. This is the short term component and its contribution is measured with $\beta_2(t)$. By setting the parameters to appropriate values, any shape of the yield curve can be constituted.

As maturity goes to zero the yield curve converges to the instantaneous spot rate, which is given by

$$\lim_{\tau \to 0} Y(t,\tau) = \beta_0(t) + \beta_1(t)$$

and the infinite yield is given by

$$\lim_{\tau \to \infty} Y(t,\tau) = \beta_0(t)$$

So unlike the cubic spline approach, the out-of-sample observations (with respect to maturity) do not diverge. The associated discount function is found after substitution of the expression for the yield curve in equation (2.19) in equation (2.2). Substitution of this discount function in equation (2.7) gives a nonlinear form, which still can be estimated by (weighted) non-linear least squares.

The third specification we consider has been introduced by Bliss (1994) and is an extended version of the Nelson-Siegel function

$$\begin{aligned} Y(t,\tau) &= \beta_0(t) + \beta_1(t)\left(\frac{1-\exp[-\tau/\gamma_1(t)]}{\tau/\gamma_1(t)}\right) + \\ &\quad +\beta_2(t)\left(\frac{1-\exp[-\tau/\gamma_2(t)]}{\tau/\gamma_2(t)} - \exp[-\tau/\gamma_2(t)]\right) \end{aligned} \tag{2.20}$$

Nelson and Siegel (1987) have also come up with this specification, but they have set $\gamma_1(t)$ equal to $\gamma_2(t)$, because for their sample of treasury bills this equation leads to overparameterization. Setting $\gamma_1(t)$ equal to $\gamma_2(t)$ in equation (2.20) results in the specification of the Nelson-Siegel function in equation (2.19). For the extended Nelson-Siegel specification the expressions for the instantaneous spot rate and the infinite yield are the same as in the Nelson-Siegel model. A second feature of the extended Nelson-Siegel model is that the alternative error term specification, given in equation (2.8), is adopted. Estimation of the extended Nelson-Siegel model proceeds similar to estimating the Nelson-Siegel model.

2.4 Data

The raw database consists of US government bond data, which are available from the CRSP bond tapes. We consider the period from January 1970 to December 1995. These data are available on a monthly basis and are the month's end quotes. The data set consists of noncallable bonds, noncallable notes, certificates of indebtness, treasury bills, callable bonds, callable notes, tax anticipation certificates of indebtness, tax anticipation bills and

others. We only take into account the noncallable bonds and notes, and the treasury bills, since all the other securities have option features in it or special tax effects, for which the pricing relation as stated in equation (2.1) does not hold.

Other studies have also eliminated other securities from the database, like notes and bonds under one year to maturity and bills under one month to maturity, since these are assumed to have special liquidity problems. An example is the filtering method that is applied to construct the Fama Bliss Discount Bond Files, see the CRSP Government Bond 1995 guide. We do not pursue this approach, since the residuals on those securities actually should be very informative in explaining these problems.

After the data reduction the data base still consists of 2248 securities and the number of securities rises from 75 in January 1970 to 225 in December 1995. There has been a little growth in the number of noncallable bonds from 20 in 1970 to 25 in 1995, treasury bill issues have decreased slightly from 35 in 1970 to 30 in 1995 and the number of treasury notes have increased from 25 in 1970 to 125 in 1995 and are therefore the mostly issued type of government bond.

For the remaining time to maturity of the bonds, it holds that the largest part has low maturity under 1 year, a respectable number of bonds are within the 1-3 year range and the 3-5 year range. The number of bonds with remaining time to maturity of 5 year and higher decreases fast. The total number of bonds which have coupon payment until ten years from the current date is still large enough to give reasonable parameter estimates in that range.

In table 2.1 the minimum, maximum and average price on bond, bills and notes are given. Especially in the period 1979-1985 bond prices have been extremely low and hence interest rates have been high. In the residual analysis we test whether other effects lead to extreme low or extreme large prices, besides from the general pricing rule.

TABLE 2.1: BOND PRICES

Year	Average	Minimum	Maximum
1970	95.23	62.13	109.28
1971	98.17	71.50	113.13
1972	98.38	76.69	110.00
1973	96.58	70.13	105.38
1974	95.76	70.50	107.31
1975	97.56	76.06	107.97
1976	99.24	80.19	111.41
1977	98.91	71.88	107.75
1978	96.67	74.19	103.44
1979	95.56	72.47	103.25
1980	93.97	66.19	118.13
1981	91.82	56.25	117.38
1982	96.93	60.44	138.56
1983	100.54	72.59	140.88
1984	97.75	66.44	130.50
1985	102.81	74.88	150.25
1986	109.48	89.13	171.66
1987	105.58	75.34	171.47
1988	103.38	78.97	159.53
1989	103.38	80.84	160.78
1990	102.84	81.84	153.66
1991	105.28	86.91	163.00
1992	107.19	89.28	164.91
1993	109.69	91.88	169.31
1994	103.76	80.06	164.63
1995	105.07	83.03	159.97

Notes: The table reports the minimum, maximum and average bond price for each year between 1970 and 1995. The data set consists of noncallable bonds, bills and notes.

2.5 Results

In our setting, it is possible to generate discount yield curves in a number of ways. We could use different approximating functions, introduce the error term in alternative ways and use different weights in the fit criterion. Here, we only consider four cases, that we think are most informative on the strength and weaknesses of alternative approaches: (a) cubic splines with spread weighted errors, (b) cubic splines with duration weighted errors, (c) Nelson-Siegel exponential functions with duration weighted errors and (d) extended Nelson-Siegel functions with duration weighted errors as specified in equation (2.8).

For each sample month the available bonds prices in that particular month are simultaneously fitted to the pricing relation. Typically, since there are more observations of bond prices than parameters to be estimated, the fit is not exact. The fit criterion is set to minimize the weighted squared errors, see equation (2.12). This results in two series. A series of fitted discount functions and a series of residuals, that represent the difference between the observed bond prices and the fitted bond prices.

For each month, the fitted discount function implies a functional representation of the yield curve. To get an impression of this yield curve, it is evaluated at maturities ranging from one month until six months and from one year until ten years. This results in a regular panel of yield curve data, that is consistent with the pricing relation for bonds in a frictionless market. In essence the data on a large cross section of bond prices are condensed to $N = 16$ observations, for $T = 300$ sample dates. Figure 2.1 gives 3D plots of the implied term structure using the Cubic Spline Method, and in figure 2.2 implied term structures from the Nelson-Siegel model are depicted. The upper part of figure 2.1 presents the implied term structure in the spread weighted cubic spline model. Typically the surface is not very smooth and a number of outliers are apparent. The bottom part presents the cubic spline model with duration weighted errors. No outliers are present, but still the surface is not extremely smooth. In figure 2.2 the duration weighted and extended Nelson-Siegel functions are depicted. Both have approximately the same levels as the duration based cubic splines, but typically the term structure is much smoother in the Nelson-Siegel models. The differences between the duration Nelson-Siegel function and the extended Nelson-Siegel function are small.

The yield curves are constructed to serve as input for empirical term structure estimation and testing. Yields and forward rates are often applied in term structure research. Keeping that in mind, we present for each method summary statistics for the yield curve, $Y_t(\tau)$, for the one-period forward rate curve, $F_{t,t+1}(\tau)$, and for the m-period forward rates of the one-month rate, $F_{t,t+m}(1)$.

The m-month forward rate curve at time t for yields with time to maturity equal to τ, is implied by the yield curve at time t, by the following relation

$$\tau F_{t,t+m}(\tau) = (\tau + m) Y_t(\tau + m) - m Y_t(m) \tag{2.21}$$

Table 2.2 gives summary statistics for the spread weighted cubic spline model, both on the constructed yield levels and on the associated one month forward rates. The tables show all kinds of misspecification. The average yield curve is not increasing and the term structure of standard deviations is not decreasing, which is normally the case. Minimum yield levels sometimes are negative or zero, which obviously makes no sense. In some cases the first order autocorrelation is low where it should be as high as 0.95 and increasing with maturity. The forward rate levels suffer from the same misspecifications as the interest rate levels. It seems that weighting the error terms with the bid-ask spread causes the outliers in the implied yield curves.

In the duration weighted spline model, we again estimate the cubic spline functions, but now on all available data in the particular month. No outliers appear and an increasing average term structure of interest rate together with a decreasing term structure of volatilities result, as shown in table 2.3. The one period autocorrelation is high and increasing with maturity, what is typically the case for yield curves. Although the implied one month forward rates do not exhibit any real outliers, they are not as smooth as the yield curves themselves.

Table 2.4 presents summary statistics for the Nelson-Siegel model with duration weighted error terms. Typically the yield curves show all the desirable properties yield curves should have, furthermore the forward rate levels show the same smooth behavior as the yield levels do. The same holds for the summary statistics of the extended Nelson-Siegel functions in table 2.5.

We also compare the m-month forward rates of the 1 month rate for the four models. The results in table 2.6 and 2.7 show that as m gets large, only the duration weighted Nelson-Siegel model produces reliable values, whereas the other models generate negative values at higher maturities. Especially in the 1980-1985 period the duration weighted Nelson-Siegel functions perform better.

To make inference on the reliability of the yield curves, we give the 95 percent confidence intervals for the time series of yields with maturities of 1 month, 1 year, 5 year and 10 year. See figure 2.3 for the spread weighted cubic spline model and figure 2.4 for the duration weighted cubic spline model.[5] The confidence interval is high for low matu-

[5]We do not give confidence intervals for the Nelson-Siegel specifications, since it is much harder to do this in a non-linear model than in a linear model as the cubic spline functions. Since the implied yield curves are very similar we expect that the confidence intervals will be similar as well.

rities and decreases with maturity. This is consistent with the findings of Fama and Bliss (1987) and Bliss (1994) who deleted short term notes and bills from the data set, because of liquidity problems. From the data section it is known that a lot of bonds with short time to maturity are present in the data set, which are exposed to liquidity problems. The confidence intervals quantify the liquidity problem. Notice that the associated discount functions show less dramatic confidence intervals, since transformation involves multiplication with time-to-maturity, which is low for short maturities. The confidence intervals are much tighter in the spread weighted spline model than in the duration weighted spline model, but this is mainly the effect of filtering some outliers out of the data, before fitting the spread spline model.[6] Also in the period 1979-1982 in which interest rates were high, the yields are less reliable then in periods of normal and low interest rates.

In practice, a lot of time series-oriented term structure models are modeled by the instantaneous spot rate. The instantaneous spot rate is not directly available, but is usually approximated by some short rate. The discount functions are also applicable for generating instantaneous spot rates, because they are implied by the discount functions. The implied instantaneous spot rates is given by the sequence $\{\alpha(t)\}_{t=1}^{T}$ for cubic spline functions and by $\{\beta_0(t)+\beta_1(t)\}_{t=1}^{T}$ for Nelson-Siegel functions. The empirical results show that the implied instantaneous spot rates are non-sensical in the case of the cubic spline model, but give a pattern that is similar to the one month spot rate in case of the Nelson-Siegel model. This implied instantaneous spot rate would be a desirable alternative to using a short rate as an approximation. The implied instantaneous spot rate is consistent with the entire yield curve. However, this implied instantaneous spot rate suffers from the same problems as the observed short rate, given the magnitude of the confidence intervals for low maturities in figure 2.4. The short rate typically consists of a lot of transitory noise which might contaminate inference about the term structure. An advantage of using the implied instantaneous rates is that the uncertainty that is available in the instantaneous spot rate is quatified in terms of a standard error, whereas in the case of the one month T-bill rate it is not.

[6] Although at first we decided not to filter any data out of our database, in fact for some cross sections the spread weighted cubic splines performed so bad, that we delete all bond prices for that day with residuals that exceeded 7 dollars in a first estimation stage. In a second stage we re-estimated the model with approximately 80 percent of the original data. Even after applying this filter a number of extreme outliers remained. The spread weighted cubic splines as proposed by McCulloch (1971) work well in a lot of cases but not always.

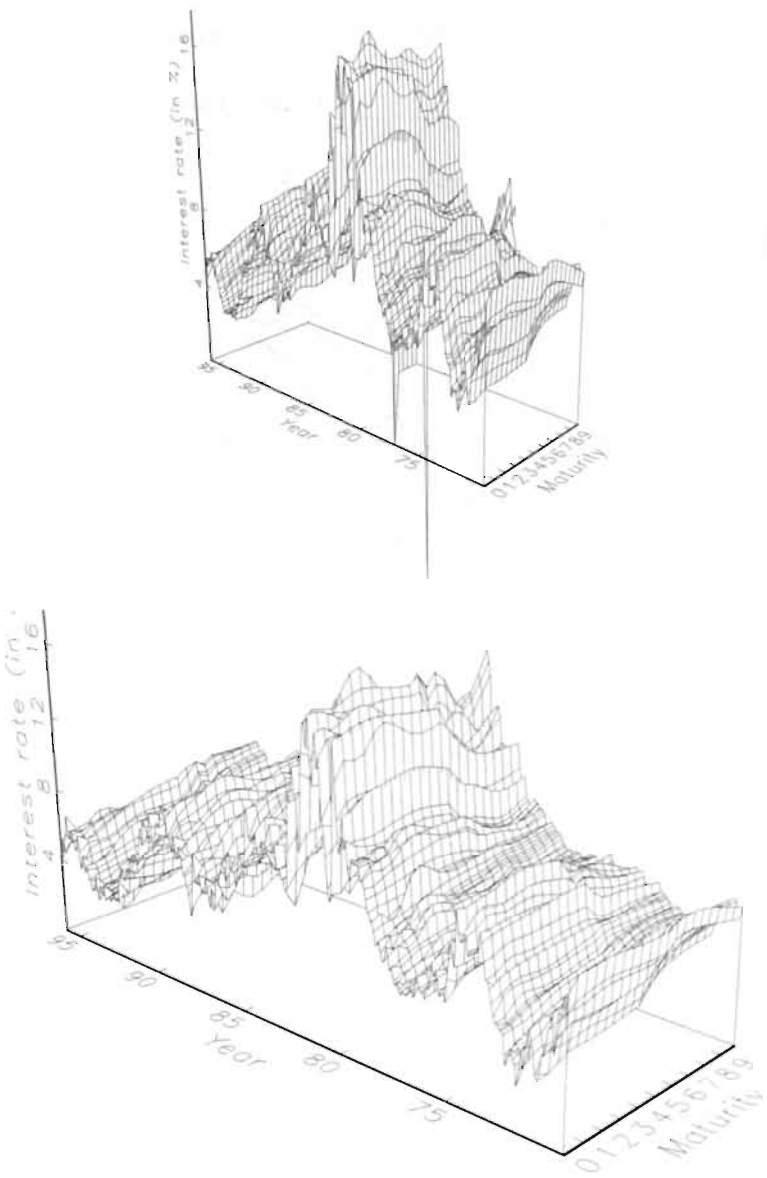

FIGURE 2.1: TERM STRUCTURES CUBIC SPLINE MODEL

The figure shows the estimated term structures for the cubic spline model with spread weigted error terms and with duration weighted error, respectively. Data consists of month end quotes between January 1970 and December 1995 on bill, notes and treasury bonds.

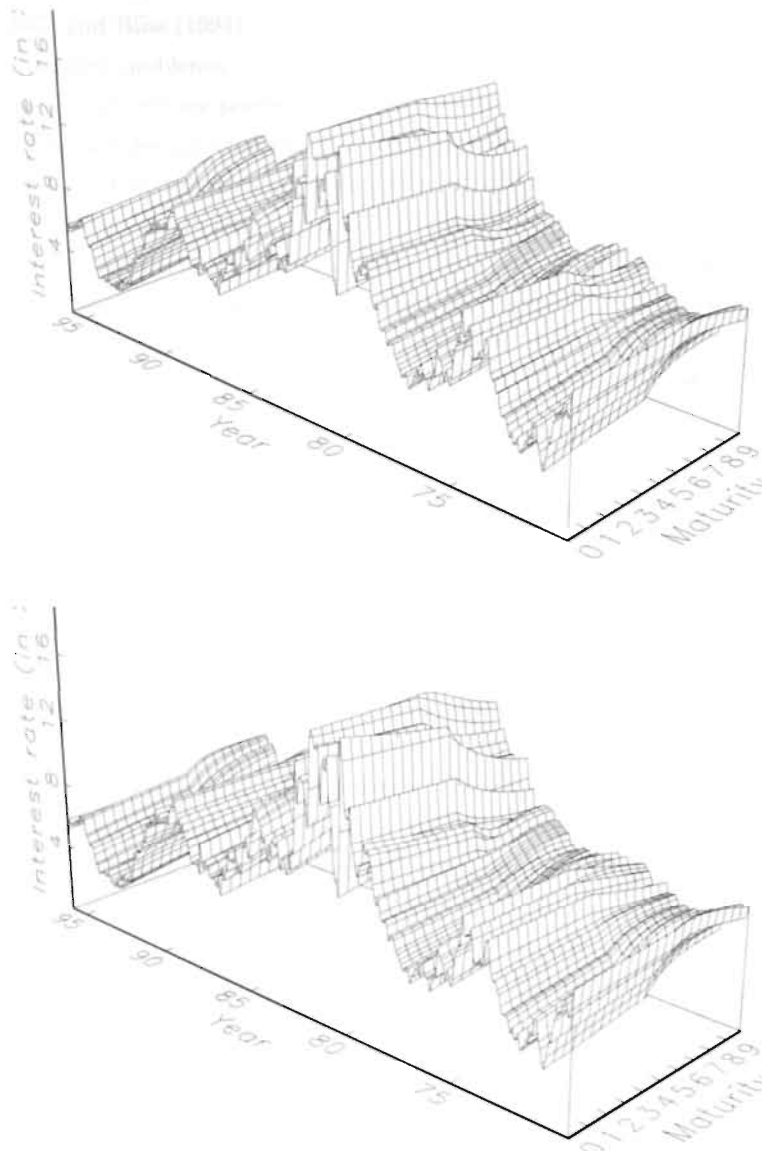

FIGURE 2.2: TERM STRUCTURES NELSON-SIEGEL MODEL

The figure shows the estimated term structures for the Nelson-Siegel model with duration weighted error terms and the extended Nelson-Siegel model, respectively. Data consists of month end quotes between January 1970 and December 1995 on bill, notes and treasury bonds.

TABLE 2.2: SUMMARY STATISTICS : SPLINE MODEL (SPREAD WEIGHTED ERRORS)

(A) Yield Levels: $Y_t(\tau)$

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	6.73	2.76	2.69	16.21	0.96
2 months	6.91	2.78	2.79	16.12	0.97
3 months	7.04	2.77	2.71	15.92	0.97
4 months	7.16	2.78	2.77	15.80	0.97
5 months	7.28	2.81	3.05	16.10	0.97
6 months	7.38	2.84	3.12	16.80	0.96
1 year	7.52	2.65	3.10	15.79	0.97
2 years	7.85	2.48	3.81	15.90	0.97
3 years	8.04	2.40	4.02	16.07	0.97
4 years	8.22	2.37	4.70	16.08	0.96
5 years	8.28	2.42	-3.86	15.78	0.85
6 years	8.39	2.36	0.00	15.55	0.90
7 years	8.47	2.36	0.00	15.48	0.90
8 years	8.57	2.39	0.00	15.64	0.90
9 years	8.66	2.41	0.00	15.90	0.66
10 years	8.70	2.76	-12.21	16.15	0.97

(B) Forward Rate Levels: $F_{t,t+1}(\tau)$

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	7.08	2.84	2.65	16.68	0.95
2 months	7.19	2.80	2.67	16.09	0.97
3 months	7.30	2.80	2.77	15.88	0.97
4 months	7.42	2.84	3.10	16.48	0.96
5 months	7.51	2.87	3.14	17.37	0.86
6 months	7.55	2.85	3.17	17.55	0.96
1 year	7.65	2.64	3.36	15.96	0.97
2 years	7.91	2.46	3.86	15.84	0.87
3 years	8.09	2.39	4.16	16.16	0.97
4 years	8.26	2.37	4.68	16.15	0.96
5 years	8.31	2.44	-5.50	15.79	0.83
6 years	8.42	2.36	-0.15	15.55	0.90
7 years	8.50	2.37	-0.13	15.50	0.90
8 years	8.60	2.39	-0.11	15.67	0.90
9 years	8.68	2.42	-0.10	15.95	0.89
10 years	8.72	2.80	-13.68	16.18	0.64

TABLE 2.3: SUMMARY STATISTICS: SPLINE MODEL (DURATION WEIGHTED ERRORS)

(A) Yield Levels: $Y_t(\tau)$

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	6.77	2.73	2.15	16.07	0.95
2 months	6.90	2.78	2.32	16.11	0.96
3 months	7.05	2.79	2.53	15.98	0.97
4 months	7.22	2.80	3.07	15.85	0.96
5 months	7.37	2.83	3.07	16.41	0.95
6 months	7.47	2.85	2.98	17.19	0.94
1 year	7.50	2.63	2.96	15.83	0.97
2 years	7.85	2.50	3.79	16.25	0.97
3 years	8.00	2.31	4.12	15.36	0.98
4 years	8.17	2.32	4.79	16.01	0.98
5 years	8.18	2.21	4.57	15.55	0.98
6 years	8.21	2.10	4.90	14.43	0.98
7 years	8.26	2.14	5.12	13.91	0.98
8 years	8.31	2.23	5.15	14.64	0.98
9 years	8.31	2.23	5.17	15.26	0.98
10 years	8.28	2.19	5.18	16.34	0.98

(B) Forward Rate Levels: $F_{t,t+1}(\tau)$

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	7.03	2.92	1.37	16.71	0.91
2 months	7.20	2.85	2.16	16.21	0.96
3 months	7.37	2.84	2.98	15.96	0.96
4 months	7.52	2.88	3.08	16.95	0.93
5 months	7.61	2.90	2.97	17.78	0.92
6 months	7.63	2.87	3.02	17.85	0.94
1 year	7.61	2.62	3.16	15.85	0.97
2 years	7.90	2.47	3.86	16.17	0.97
3 years	8.05	2.31	4.21	15.42	0.98
4 years	8.20	2.31	4.85	16.07	0.98
5 years	8.20	2.19	4.59	15.49	0.98
6 years	8.23	2.09	4.98	14.35	0.98
7 years	8.29	2.15	5.15	14.05	0.98
8 years	8.32	2.23	5.17	14.70	0.98
9 years	8.32	2.23	5.19	15.36	0.98
10 years	8.29	2.19	5.20	16.43	0.98

TABLE 2.4: SUMMARY STATISTICS: NELSON-SIEGEL METHOD (DURATION WEIGHTED ERRORS)

(A) Yield Levels: $Y_t(\tau)$

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	7.20	2.84	2.73	16.18	0.97
2 months	7.23	2.82	2.82	16.16	0.97
3 months	7.26	2.80	2.91	16.14	0.97
4 months	7.30	2.78	3.00	16.12	0.97
5 months	7.33	2.76	3.08	16.09	0.97
6 months	7.36	2.74	3.14	16.07	0.97
1 year	7.52	2.64	3.41	16.03	0.97
2 years	7.80	2.50	3.88	15.96	0.97
3 years	8.02	2.40	4.28	15.89	0.98
4 years	8.19	2.34	4.62	15.83	0.98
5 years	8.33	2.30	4.91	15.77	0.98
6 years	8.43	2.28	5.08	15.72	0.98
7 years	8.52	2.28	5.15	15.68	0.98
8 years	8.59	2.28	5.19	15.64	0.98
9 years	8.64	2.29	5.21	15.60	0.98
10 years	8.69	2.31	5.20	15.57	0.99

(B) Forward Rate Levels: $F_{t,t+1}(\tau)$

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	7.27	2.80	2.91	16.14	0.97
2 months	7.30	2.78	3.01	16.12	0.97
3 months	7.33	2.76	3.08	16.09	0.97
4 months	7.36	2.74	3.14	16.07	0.97
5 months	7.39	2.72	3.18	16.07	0.97
6 months	7.42	2.71	3.23	16.06	0.97
1 year	7.58	2.61	3.49	16.02	0.97
2 years	7.85	2.48	3.96	15.95	0.98
3 years	8.06	2.39	4.35	15.88	0.98
4 years	8.22	2.33	4.69	15.82	0.98
5 years	8.35	2.30	4.96	15.76	0.98
6 years	8.46	2.28	5.11	15.71	0.98
7 years	8.54	2.28	5.18	15.67	0.98
8 years	8.61	2.28	5.21	15.63	0.98
9 years	8.66	2.29	5.22	15.59	0.98
10 years	8.70	2.31	5.22	15.56	0.99

TABLE 2.5: SUMMARY STATISTICS: EXTENDED NELSON-SIEGEL METHOD

(A) Yield Levels: $Y_t(\tau)$

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	7.17	2.86	2.70	16.22	0.96
2 months	7.21	2.84	2.81	16.12	0.96
3 months	7.25	2.81	2.91	16.12	0.96
4 months	7.28	2.79	2.97	16.11	0.97
5 months	7.32	2.77	3.03	16.10	0.97
6 months	7.36	2.75	3.09	16.09	0.97
1 year	7.54	2.65	3.43	16.08	0.97
2 years	7.83	2.50	3.86	15.10	0.98
3 years	8.02	2.38	4.22	15.92	0.98
4 years	8.15	2.28	4.54	15.67	0.98
5 years	8.24	2.20	4.82	15.39	0.98
6 years	8.29	2.14	5.05	14.11	0.98
7 years	8.33	2.09	5.13	14.85	0.98
8 years	8.34	2.06	5.16	14.62	0.98
9 years	8.33	2.04	5.17	14.42	0.98
10 years	8.31	2.04	5.16	14.24	0.98

(B) Forward Rate Levels: $F_{t,t+1}(\tau)$

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	7.25	2.81	2.91	16.12	0.96
2 months	7.29	2.79	2.97	16.11	0.97
3 months	7.32	2.77	3.04	16.10	0.97
4 months	7.36	2.75	3.09	16.10	0.97
5 months	7.39	2.73	3.15	16.09	0.97
6 months	7.43	2.72	3.21	16.07	0.97
1 year	7.61	2.62	3.53	16.12	0.97
2 years	7.87	2.48	3.93	16.10	0.98
3 years	8.06	2.36	4.28	15.91	0.98
4 years	8.18	2.27	4.60	15.64	0.98
5 years	8.26	2.19	4.86	15.36	0.98
6 years	8.31	2.13	5.09	15.08	0.98
7 years	8.34	2.08	5.16	14.82	0.98
8 years	8.35	2.05	5.18	14.59	0.98
9 years	8.34	2.04	5.19	14.39	0.98
10 years	8.32	2.03	5.18	14.21	0.98

TABLE 2.6: EVOLUTION OF FORWARD RATE IN SPLINE MODELS

(A) Spread weighted: $F_{t,t+m}(1)$

m	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	6.73	2.76	2.69	16.21	0.96
2 months	7.08	2.84	2.65	16.68	0.95
3 months	7.29	2.80	2.56	15.70	0.95
4 months	7.53	2.86	2.78	16.82	0.94
5 months	7.78	3.06	3.16	19.62	0.88
6 months	7.86	3.04	3.20	20.91	0.90
1 year	7.97	2.60	3.98	16.80	0.93
2 years	8.02	2.26	3.37	19.73	0.82
3 years	8.65	2.41	4.25	17.59	0.95
4 years	8.80	2.76	2.46	31.48	0.79
5 years	8.39	8.94	-142.80	15.49	0.02
6 years	8.91	2.66	0.00	15.30	0.86
7 years	9.07	2.78	0.00	16.64	0.86
8 years	9.39	3.09	0.00	23.09	0.81
9 years	9.07	9.71	-147.52	19.07	0.09
10 years	9.48	11.35	-175.75	27.28	0.08

(B) Duration weighted: $F_{t,t+m}(1)$

m	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	6.77	2.73	2.15	16.07	0.95
2 months	7.03	2.92	1.37	16.71	0.91
3 months	7.36	2.87	2.89	15.83	0.94
4 months	7.73	3.02	3.87	18.20	0.82
5 months	7.97	3.23	2.75	20.61	0.72
6 months	7.94	3.08	2.52	21.09	0.84
1 year	7.76	2.54	3.16	15.97	0.93
2 years	7.98	2.13	2.97	15.03	0.85
3 years	8.67	2.42	4.85	16.01	0.96
4 years	8.45	2.44	3.45	17.83	0.94
5 years	8.17	1.81	3.96	15.90	0.91
6 years	8.51	2.48	5.43	17.09	0.95
7 years	8.66	3.24	5.09	21.71	0.95
8 years	8.51	3.22	4.01	20.07	0.92
9 years	8.16	3.63	-1.07	25.15	0.93
10 years	7.90	3.80	-3.06	25.80	0.92

TABLE 2.7: EVOLUTION OF FORWARD RATE IN NELSON-SIEGEL MODELS

(A) Duration weighted: $F_{t,t+m}(1)$

m	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	7.20	2.84	2.73	16.18	0.97
2 months	7.27	2.80	2.91	16.14	0.97
3 months	7.33	2.76	3.08	16.09	0.97
4 months	7.39	2.72	3.18	16.07	0.97
5 months	7.45	2.69	3.28	16.05	0.97
6 months	7.51	2.65	3.35	16.04	0.97
1 year	7.82	2.50	3.80	15.96	0.97
2 years	8.28	2.33	4.68	15.82	0.98
3 years	8.59	2.26	5.29	15.70	0.98
4 years	8.79	2.25	5.59	15.60	0.98
5 years	8.92	2.26	5.65	15.51	0.98
6 years	9.00	2.30	5.62	15.44	0.98
7 years	9.05	2.35	5.37	15.50	0.99
8 years	9.08	2.42	5.07	15.58	0.99
9 years	9.09	2.51	4.76	15.64	0.99
10 years	9.09	2.63	4.46	15.65	0.98

(B) Extended: $F_{t,t+m}(1)$

m	Average	Standard Deviation	Minimum	Maximum	Auto-correlation
1 month	7.23	2.88	2.69	16.25	0.96
2 months	7.42	2.90	2.88	16.06	0.96
3 months	7.61	2.95	3.04	16.57	0.96
4 months	7.79	3.03	3.17	18.50	0.96
5 months	7.96	3.13	3.29	20.30	0.96
6 months	8.11	3.24	3.40	21.98	0.96
1 year	8.89	3.95	3.91	29.71	0.95
2 years	9.73	4.69	4.61	35.78	0.95
3 years	9.91	4.44	5.10	33.64	0.94
4 years	9.69	3.61	5.47	26.93	0.95
5 years	9.23	2.64	5.60	18.77	0.95
6 years	8.67	2.16	5.53	14.33	0.97
7 years	8.06	2.66	-1.28	14.14	0.97
8 years	7.46	3.72	-10.20	13.93	0.96
9 years	6.89	4.89	-18.22	13.74	0.96
10 years	6.37	6.01	-25.26	13.60	0.96

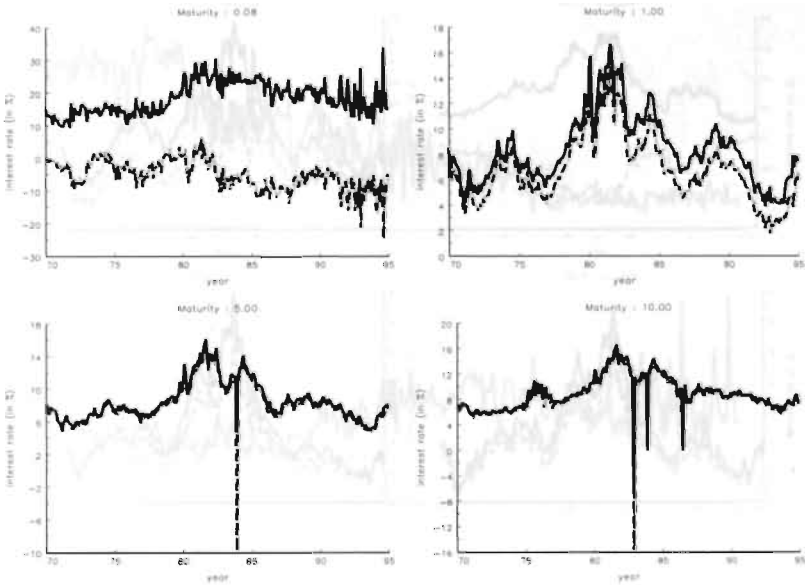

FIGURE 2.3: TIME SERIES OF INTEREST RATES CONFIDENCE INTERVALS

Each picture presents the 95% confidence interval for interest rates with maturities of 1 month, 1 year, 5 years and 10 years. Data consists of month end quotes between January 1970 and December 1995 on bill, notes and treasury bonds. The yield curves have been estimated using the spread weighted cubic spline method.

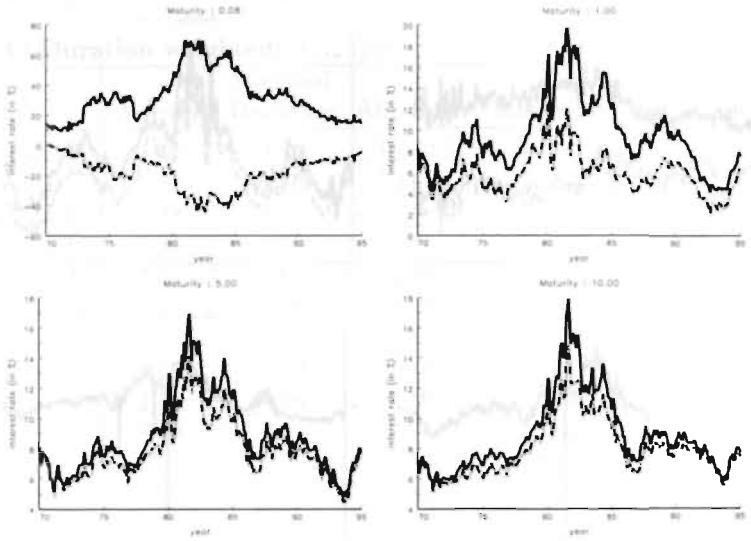

FIGURE 2.4: TIME SERIES OF INTEREST RATES CONFIDENCE INTERVALS

Each picture presents the 95% confidence interval for interest rates with maturities from 1 month to 10 years. Data consists of end months quotes between January 1970 and December 1995 on bill, notes and treasury bonds. The yield curves have been estimated using the duration weighted cubic spline method.

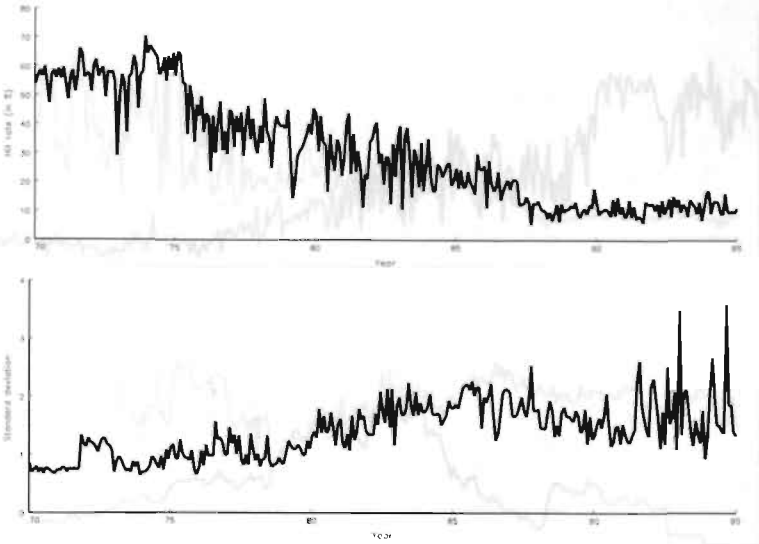

FIGURE 2.5: FIT SPREAD WEIGHTED CUBIC SPLINE METHOD

In the upper figure the hit rate is depicted, where the hit rate is the number of residuals that is within the bid-ask spread. The lower figure shows the standard deviation of the residuals for each cross section estimate. Data consists of month end quotes between January 1970 and December 1995 on bill, notes and treasury bonds. The yield curves have been estimated using the cubic spline method with spread weighted error terms.

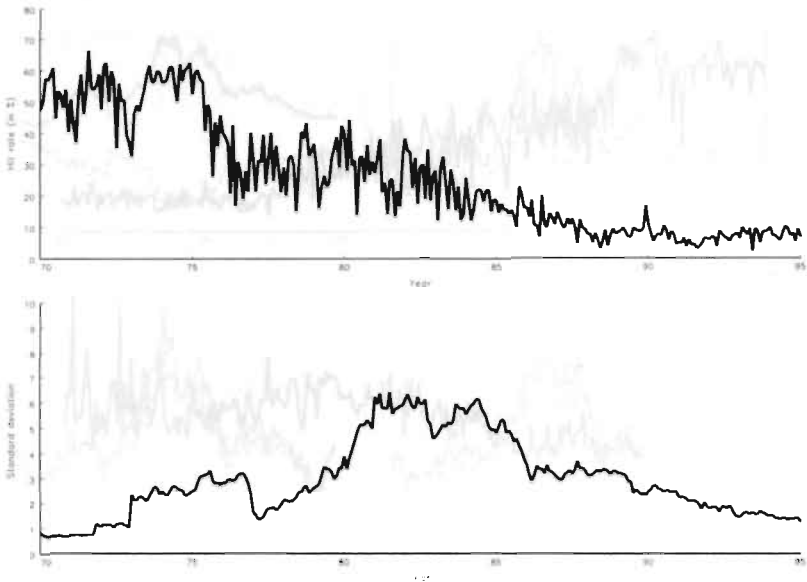

FIGURE 2.6: FIT DURATION WEIGHTED CUBIC SPLINE MODEL

In the upper figure the hit rate is depicted, where the hit rate is the number of residuals that is within the bid-ask spread. The lower figure shows the standard deviation of the residuals for each cross section estimate. Data consists of month end quotes between January 1970 and December 1995 on bill, notes and treasury bonds. The yield curves have been estimated using the cubic spline method with duration weighted error terms.

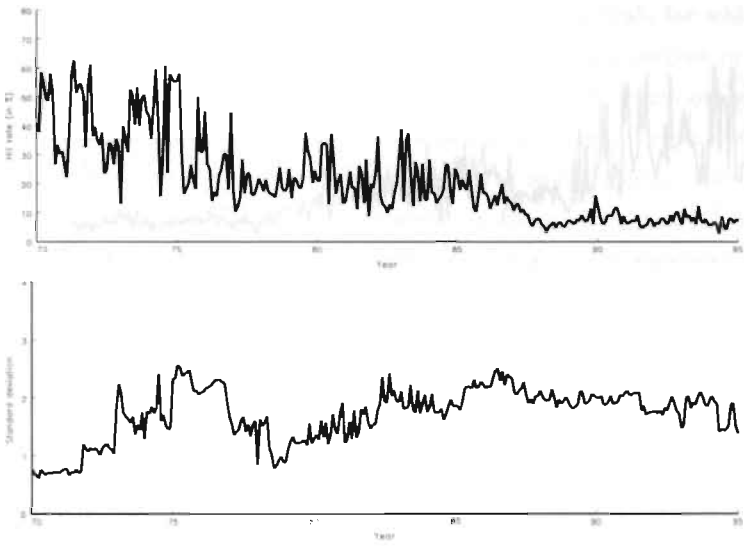

FIGURE 2.7: FIT DURATION WEIGHTED NELSON-SIEGEL MODEL

In the upper figure the hit rate is depicted, where the hit rate is the number of residuals that is within the bid-ask spread. The lower figure shows the standard deviation of the residuals for each cross section estimate. Data consists of month end quotes between January 1970 and December 1995 on bill, notes and treasury bonds. The yield curves have been estimated using the Nelson-Siegel method with duration weighted error terms.

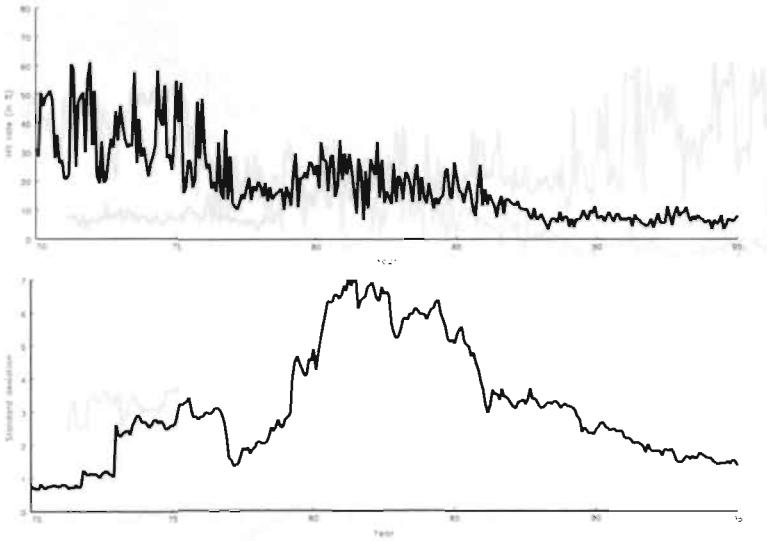

FIGURE 2.8: FIT DURATION WEIGHTED EXTENDED NELSON-SIEGEL MODEL

In the upper figure the hit rate is depicted, where the hit rate is the number of residuals that is within the bid-ask spread. The lower figure shows the standard deviation of the residuals for each cross section estimate. Data consists of month end quotes between January 1970 and December 1995 on bill, notes and treasury bonds. The yield curves have been estimated using the extended Nelson-Siegel method with duration weighted error terms.

The residual analysis consists of two parts. The first part provides a measure of fit for each cross section. Two measures are applied, the hit rate and the standard deviation of the residuals. The hit rate gives the percentage of all residuals for which the model price is within the bid-ask spread. The hit rate is as high as 60 percent in the seventies and decreases in time to about 10 percent in 1995, see the top parts of figures 2.5, 2.6, 2.7 and 2.8. This pattern is consistent for the four approaches. An explanation is that the number of bonds, which have to be priced by the same relation, increases in time. The standard deviation of residuals are presented in the bottom parts of figure 2.5, 2.6, 2.7 and 2.8. In the duration weighted spline model and in the extended Nelson-Siegel specification, especially in the period 1980-1985, the standard deviation is much higher than in the duration weighted Nelson-Siegel specification. It seems that a period with high interest rates or a change of regime has a greater influence on the former two than on the latter one.

The associated residuals from the fit of the raw bond data to the pricing relation provides information on all the other variables that effect the price of a bond. This set of residuals is explored to detect which other variables influence the price of a bond, or to investigate which type of bonds is highly effected by these other variables or under what circumstances (e.g. in what time periods). These variables could be interpreted as omitted pricing factors. Notice that we only consider linear relations between residuals and pricing factors, while in general these factors might influence the price of a bond in a non-linear way. For example tax effects deal with timing of cash outflows. If an effect shows up in the same way in each case then we might assume that the associated factor indeed has been omitted. In table 2.8 we report the regression results of the remaining error terms to a number of explanatory variables. We have performed one large regression that includes the residuals of the bond prices in all months simultaneously. Following Bliss (1994), we include as explanatory variables premium, i.e. the amount the bond price is above par, discount, i.e. the amount the bond price is below par, and maturity. Furthermore, dummy variables are included for the type of issue, to detect whether there is systematically mispricing under bonds (type 1), notes (type 2) or bills (type 3).

Also a time dummy is included, which is set to one whenever we observed data in the period 1979 - 1985, and which is zero otherwise. Apparently a regime shift has taken place in this period, leading to interest rates that are significantly higher than in other periods. Bliss (1994) states that particularly in the early 1980's, the extremely high interest rates produced heterogenous samples of deep discount and high premium bonds with the distributions shifting with levels of interest rates. The extreme values would produce perceived price distortions under the PV-only pricing assumption if either tax-clientele or

tax-timing were important.

In table 2.8 particularly the duration weighted cubic spline model and the extended Nelson-Siegel model show similar effects in the residuals. The duration weighted Nelson-Siegel model sometimes is consistent with the former two. The spread weighted cubic spline model produces different regression results, but there are at least two reasons why they would not be the same. First of all the model was not estimated on the full sample, but outliers have been filtered out of the data set. Secondly, the spread weighted cubic spline model still exhibits a number of extreme outliers, which may contaminate the results.

All models agree upon the effect of premium or discount effects to the error terms. Premiums are negatively related to the error term, whereas discounts are positively related. Not all models are unanimous with respect to the time to maturity, although three of them expect to find a significantly negative relation between the level of the error term and the time to maturity. The effect of the type of issue shows that bonds are positively related, notes are positively related (although the duration weighted Nelson-Siegel model suggests the opposite), bills are negatively related. All models agree upon the effect in the period 1979-1985, there is a negative effect on the error terms.

TABLE 2.8: ERROR TERM ANALYSIS

Explanatory Variables	Spread Spline	Duration Spline	Duration Nelson	Extended Nelson
Premium	-0.0048	-0.0487	-0.0186	-0.0641
	(0.0013)	(0.0019)	(0.0014)	(0.0021)
Discount	0.0372	0.1164	0.0544	0.1255
	(0.0027)	(0.0039)	(0.0028)	(0.0043)
Maturity	-0.0204	-0.0638	0.0471	-0.0448
	(0.0022)	(0.0033)	(0.0023)	(0.0036)
Type 1	0.4686	2.3367	0.0168	2.3687
	(0.0419)	(0.0615)	(0.0443)	(0.0681)
Type 2	0.0204	0.1615	-0.1220	0.1648
	(0.0153)	(0.0225)	(0.0162)	(0.0249)
Type 3	-0.0640	-0.1729	-0.0578	-0.1889
	(0.0223)	(0.0327)	(0.0236)	(0.0362)
period 79-85	-0.0242	-0.2117	-0.1749	-0.2032
	(0.0230)	(0.0338)	(0.0244)	(0.0374)

Notes: The table reports the regression coefficients to explain the errors from the pricing relation to the bond prices, standard errors are within parentheses. A premium, discount and maturity effect are included. Furthermore dummies for the type of issue (bond, note or bill) and a dummy for the period 1979-1985 are included.

2.6 Concluding remarks

In this chapter we have compared four models which have been proposed in the literature to extract discount yield data from coupon bearing bond data. We have compared the models to each other and find that duration weighted methods show less outliers than spread weighted models. The exponential specifications as proposed by Nelson and Siegel (1987) give smoother curves than the cubic spline models, but the differences are typically small.

In chapters 3, 4 and 5, we concentrate on the cross sectional dimension of yield curve models. We employ panel data techniques to estimate and to test various term structure models. The use of the regular panel as constructed by the statistical methods in this chapter, simplify the econometric analysis considerably. In this way we avoid the non-linearities, that stem from the fact that bonds normally carry coupon-payments. The confidence intervals around the constructed yield curves show that in particular at the very short end of the yield curve the data is less reliable and term structure models may perform poorly. Also in the period 1979-1982, when interest rates are high, the confidence intervals are wider and hence we would expect term structure models to give less accurate results.

The smoothness of the exponential Nelson-Siegel functions also results in smoother forward rate curves and in instantaneous spot rates that show no outliers, in contrast with the cubic spline functions. The duration weighted Nelson-Siegel functions are virtually insensitive to different regimes, whereas the extended Nelson-Siegel method and the duration weighted cubic spline method are. The implied forward rate curves, for the one month rate for more than one period ahead, show outliers in all models, except for the duration weighted Nelson-Siegel model. Empirical tests of, for example, the expectations hypothesis of interest rates use these forward rate, see for example Evans and Lewis (1994). For testing this type of models duration weighted Nelson-Siegel yield curves seems more appropriate. The use of constructed forward rates in tests of term structure models however also incorporate the danger of smoothing away too much information available in the raw data. In chapter 3 forward rates are used in the empirical analysis of the expectations hypothesis. Since the Nelson-Siegel functions are based upon the expectations hypothesis, we prefer to use yields and forwards that are constructed using the cubic spline method.

In the remaining of this thesis we will only deal with models that satisfy the complete market assumptions and also all sorts of frictions are not included. The panel of yield curves as constructed has only taken account of this part of the data.

Chapter 3

The Time Series and Cross Sectional Properties of Interest Rates Risk Premia

3.1 Introduction

The expectations hypothesis of the term structure of interest rates has been studied in the literature for a long time. The early literature[1] deals with the *pure* expectations hypothesis, which is presented in different versions. The first version states that the excess holding period return on a bond is equal for all maturities. The second version of the pure expectations hypothesis states that the yield on a long term bond is a weighted average of current and future expected short rates over the life of the yield of the long term bond. In the third version, the forward rate is claimed to be an unbiased predictor of the future spot rate.

Although Cox, Ingersoll and Ross (1981) show that these three variants are logically incompatible, Campbell (1986) demonstrates that they are not substantively dissimilar, as they are well approximated by a family of linear approximations which is internally consistent. In periods of high and volatile rates and for very long-term bonds care must be taken in using the linearization.

The very stringent hypothesis posed in either of the three variants appears not to hold and the inclusion of a constant risk premium has been adopted. Hicks (1939) postulated the liquidity premium theory, which states that borrowers of long term capital are required to increase their promised return to compensate the preferred liquidity of lenders. Hence the longer the time to maturity of a bond, the greater the liquidity premium would need

[1] For a thorough overview of the literature on the expectations hypothesis we refer to Melino (1988).

The current term structure of interest rates also provides a forward rate curve. The forward rate at time t is the rate at which investors currently agree to enter into a contract at time $t+1$, that has maturity τ. In a world of certainty, the current forward rate is equal to the expected future yield. The one-period forward rate, $F_{t,t+1}(\tau)$, is given by

$$(\tau - 1)F_{t,t+1}(\tau - 1) = \tau Y_t(\tau) - Y_t(1) \tag{3.3}$$

This definition of forward rates follows from two investment strategies that yield the same result. In the first strategy an investor decides to enter into a contract that holds for τ periods. In the second strategy, the investor decides to invest for one period and to invest the payoff of this investment for $\tau - 1$ periods, starting one period from now. Because both strategies start now and end at the same time, in the absence of arbitrage opportunities, their returns will be equal. Clearly, the forward rate on the left-hand-side of equation (3.3) is implied by current yields on the right-hand-side.

Consider an investor who faces at time t the possibility to invest in alternative discount bonds. Investing in a discount bond that has a remaining time-to-maturity of one period is riskless, since at maturity the principal is repaid. The one-period return on this discount bond is the one-period yield, $Y_t(1)$. The one-period return on a general discount bond is the holding period return. We define the value of the holding period return in excess of the return on a risk free investment as the excess holding period return

$$\zeta_{t+1}(\tau) = H_{t,t+1}(\tau) - Y_t(1) \tag{3.4}$$

The excess holding period return denotes the additional return on a risky investment over a certain investment for the same period. An interesting issue is whether this excess holding period return is zero on average (the pure expectations hypothesis), or positive (the expectations hypothesis). In the latter case the investor is rewarded with a higher expected return to compensate for the additional amount of risk. We refer to the expected excess holding period return as the risk premium, which quantifies the expected excess return over the riskfree rate. The risk premium, $\psi_t(\tau)$, is defined as

$$\psi_t(\tau) = E_t[\zeta_{t+1}(\tau)] = E_t[H_{t,t+1}(\tau)] - Y_t(1) \tag{3.5}$$

We focus on the properties of this risk premium. Using the relations in equations (3.1), (3.2) and (3.3) the risk premium as defined in equation (3.5) can be rewritten as

$$\psi_t(\tau) = (\tau - 1)[F_{t,t+1}(\tau - 1) - E_t(Y_{t+1}(\tau - 1))] \tag{3.6}$$

This results in a second interpretation of the risk premium, which is the difference between the forward rate and the expected future yield, premultiplied with a maturity dependent

factor. Under the pure expectations hypothesis the forward rate is assumed to be equal to the expected future yield and hence the risk premium is zero in that case. From equation (3.6) a testable expression of the expectations hypothesis follows, which links the future yield to the current forward rate

$$\tau Y_{t+1}(\tau) = \tau F_{t,t+1}(\tau) - \psi_t(\tau+1) + \tau\left[Y_{t+1}(\tau) - E_t\left(Y_{t+1}(\tau)\right)\right] \tag{3.7}$$

The future yield consists of three components: the current forward rate, a risk premium and an unexpected excess yield. This expression encompasses the pure expectation hypothesis, when $\psi_t(\tau+1) = 0$, and the expectations hypothesis, otherwise. The expected value of future interest rate is the current forward rate plus a risk premium. The relation in (3.7) is the fundamental relation that is our starting point for the empirical analysis. Note that we choose to run levels regressions because subtracting either yields or forward rates on both sides to render the regression variables "more stationary", as was suggested by Campbell and Shiller (1987), increases the potential for biases in small samples, see Bekaert, Hodrick and Marshall (1997).

3.3 The econometric framework

In the literature, testing the pure expectations hypothesis typically boils down to estimating the regression model

$$\tau Y_{t+1}(\tau) = \alpha + \beta\tau F_{t,t+1}(\tau) + \varepsilon_{t+1}(\tau) \tag{3.8}$$

where α and β are unknown parameters. Testing for the pure expectations hypothesis amounts to testing the hypothesis that $\alpha = 0$ and $\beta = 1$. The error term $\varepsilon_{t+1}(\tau)$ is given by

$$\varepsilon_{t+1}(\tau) = -\psi_t(\tau+1) + \tau\left[Y_{t+1}(\tau) - E_t\left(Y_{t+1}(\tau)\right)\right] \tag{3.9}$$

which implies that the error term consists of two components. One depends upon the risk premium, the other one is the unexpected excess yield. Leaving the risk premium in the error term does not allow us to gain insight into its properties, apart from its average value, that is measured through the estimate for α. In our view knowledge about the properties of the risk premium is important. It not only enables a better understanding of pricing of risk in the economy, but also it may be able to explain *why* the expectations hypothesis is rejected. Because both the risk premium and the unexpected excess yield are unobservable, it is in general not clear to which of the two we should attribute an effect in the error term as given in equation (3.9). Once we are able to discern the risk premium component from the unexpected excess yield component, the regression results could actually provide clear

insight as to which component can account for the rejection of the expectations hypothesis, which is fairly common in the literature, see Melino (1988) for an overview. Panel data techniques are a convenient tool to disentangle risk premia and unexpected excess yields.

Next, we set up the panel data framework for the expectations hypothesis regression model of equations (3.8) and (3.9). For each time-to-maturity the pure expectations hypothesis states that $\alpha = 0$ and $\beta = 1$, hence pooling the data for different maturities gives the following model

$$\begin{bmatrix} \tau_1 Y_{t+1}(\tau_1) \\ \vdots \\ \tau_N Y_{t+1}(\tau_N) \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \alpha + \begin{bmatrix} \tau_1 F_{t,t+1}(\tau_1) \\ \vdots \\ \tau_N F_{t,t+1}(\tau_N) \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_{t+1}(\tau_1) \\ \vdots \\ \varepsilon_{t+1}(\tau_N) \end{bmatrix} \tag{3.10}$$

or in matrix notation

$$Y_{t+1} = \iota\alpha + F_{t,t+1}\beta + \varepsilon_{t+1} \qquad t = 1, \ldots, T \tag{3.11}$$

where Y_{t+1} is the $(N \times 1)$ vector of maturities times yields at time $t+1$ for maturities $\tau_1, \ldots, \tau_N$, ι is the $(N \times 1)$ vector of ones, $F_{t,t+1}$ is the $(N \times 1)$ vector of maturities times forward rates at time t, and ε_{t+1} denotes the $(N \times 1)$ vector of error terms which consists of a risk premium and the unexpected excess yield. Stacking the vectors for all time observations $t = 1, \ldots, T$ results in the panel data model. We refer to this model as the pooled case. In this specification we do not explicitly account for the existence of the risk premium, but relegate it to the error term. Testing this specification results in more efficient parameter estimates than in the univariate time series models. The main motivation for considering the pooled case is because it may adress bias and dispersion in the small sample distributions of the parameter estimates, which were documented by Bekaert, Hodrick and Marshall (1997). The biases arise because of the extreme persistence in short rates, which implies that the asymptotic distributions cannot be relied upon. In the pooled case all yields in a cross section are incorporated simultaneously to estimate the parameters. This leads to more efficient estimates and because of the increased number of observation the small sample bias becomes less important.

In the following, this model specification is extended to incorporate explicitly the risk premia as incidental variables. In a panel data framework the usual approach to model incidental variables is either as fixed effects, including a dummy for each observation, or as random effects by specifying a stochastic process for the variables in question. Direct application of either approach, however, is not possible in our model. In order to take account of the risk premium we extract it from the error term and rewrite the model as

$$\begin{bmatrix} \tau_1 Y_{t+1}(\tau_1) \\ \vdots \\ \tau_N Y_{t+1}(\tau_N) \end{bmatrix} = \begin{bmatrix} \tau_1 F_{t,t+1}(\tau_1) \\ \vdots \\ \tau_N F_{t,t+1}(\tau_N) \end{bmatrix} \beta - \begin{bmatrix} \psi_t(\tau_1) \\ \vdots \\ \psi_t(\tau_N) \end{bmatrix} + \begin{bmatrix} \mathrm{UEY}_{t+1}(\tau_1) \\ \vdots \\ \mathrm{UEY}_{t+1}(\tau_N) \end{bmatrix} \tag{3.12}$$

where we introduce

$$\mathrm{UEY}_{t+1}(\tau) \equiv \tau \left(Y_{t+1}(\tau) - E_t[Y_{t+1}(\tau)] \right) \quad \tau = \tau_1, ..., \tau_N \tag{3.13}$$

for the unexpected excess yield. In matrix notation it holds that

$$Y_{t+1} = F_{t,t+1}\beta - \psi_t + \mathrm{UEY}_{t+1} \qquad t = 1, \ldots, T \tag{3.14}$$

Note that we have split the error term from equation (3.9) into two parts, the risk premium ψ_t and the unexpected excess yield UEY_{t+1}. The unexpected excess yield is treated as an error term that takes account of the cross sectional relations between the univariate series, through the adoption of a covariance matrix specification. We adopt the following specification for the error terms

$$\mathrm{UEY}_{t+1} \sim N(0, \Sigma)$$

where we specify the exact form of the covariance matrix below.

Our goal is to model the risk premium explicitly, where the fixed maturity effects estimator seems a natural candidate to learn about its maturity properties. The fixed maturity effects estimator treats the risk premium as a constant in time and assumes that it is only different whenever the maturity of the yield is different, hence

$$\psi_1(\tau) = \cdots = \psi_T(\tau) \qquad \tau = \tau_1, ..., \tau_N \tag{3.15}$$

This estimator quantifies the relative levels of risk premia for yields with different time-to-maturity. The pure expectations hypothesis postulates that they are zero for all maturities. The liquidity preference theory suggests that the risk premium increases with time to maturity. In the appendix we derive the panel data estimator[2] for the fixed effects, the associated estimator for β, and expressions for standard errors. This results in the following estimator for the fixed maturity effects

$$\begin{bmatrix} \hat{\psi}(\tau_1) \\ \vdots \\ \hat{\psi}(\tau_N) \end{bmatrix} = \frac{1}{T} \sum_{t=1}^{T} (Y_t - F_{t,t+1}\beta) - \frac{1}{T} \sum_{t=1}^{T} \mathrm{UEY}_t \tag{3.16}$$

[2]See Baltagi (1995) and Hsiao (1986) for textbook treatments of panel data models.

We split the estimator into an observed part and an unobserved part. The unobserved part represents the sample average of the unexpected excess return, which approaches zero by the law of large numbers. The fixed maturity effects estimator is equal to the estimator we find when the unexpected excess yield would not be incorporated in the model, and hence this estimator specifically describes the behavior of the risk premium in the maturity dimension.

The fixed maturity effects specification is interesting in itself, but at the same time somewhat restrictive: the risk premia are allowed to vary only with maturity and not with time. In what follows, we introduce time variation into the model by employing ingredients from a time effects panel model. The use of a random time effects model lies in *combining* it with the fixed maturity effects model in order to incorporate both the maturity and time dimensions appropriately.

Using equation (4) and the fact that $\psi_t(\tau) = E_t\zeta_{t+1}(\tau)$, we derive

$$Y_{t+1} = F_{t,t+1}\beta - \zeta_{t+1} + \mathrm{UEY}_{t+1} + \mathrm{UEH}_{t+1} \tag{3.17}$$

where the term given by

$$\mathrm{UEH}_{t+1}(\tau) = [\zeta_{t+1}(\tau) - E_t\zeta_{t+1}(\tau)] \qquad \tau = \tau_1, \ldots, \tau_N \tag{3.18}$$

denotes the unexpected excess holding period return. Obviously, the right-hand-side now consists of three terms which are indiscernible: the excess holding period return, ζ_{t+1}, the unexpected excess yield, UEY_{t+1}, and the unexpected excess holding period return, UEH_{t+1}. The last two terms are unexpected effects and are easily eliminated by taking conditional expectations. We impose an AR(1) process for the sum of excess holding period return, unexpected excess yield and unexpected excess return, denoted by

$$\lambda_{t+1} = -\zeta_{t+1} + \mathrm{UEY}_{t+1} + \mathrm{UEH}_{t+1} \tag{3.19}$$

This reduces the number of parameters to the number of unknown parameters in an AR(1) process which is only three. Note that taking conditional expectations in the estimated time series process results in a time series of risk premia

$$E_t\lambda_{t+1} = -E_t\zeta_{t+1} = -\psi_t \qquad t = 1, \ldots, T-1 \tag{3.20}$$

since the conditional expectations of the second and third components are zero. Because we are only interested in the risk premia, the modeling decision relies heavily on the assumptions we make for the excess holding period returns, whereas the second and third components are not relevant since they are eliminated from the equation by taking conditional expectations. The risk premium is allowed to be different from period to period

with the arrival of news or with changes in the economic regime. The AR(1) process allows for such changes and provides us with a measure of persistence of the risk premium over time.

We assume that the excess holding period returns for yields with different time to maturity are driven by the same dynamics: we impose a time-invariant one-factor model in time dimension. We also incorporate the maturity dimension of risk premia. The excess holding period returns of the long term yield serves as a base case. The risk premia of all other yields are assumed to be related to this process through a scaling factor

$$\zeta_{t+1}(\tau) = Z(\tau)\zeta_{t+1} \qquad \tau = \tau_1, ..., \tau_N \tag{3.21}$$

with $Z(\tau_N) = 1$. Note that the scaling factors $Z(\tau_1)$,..., $Z(\tau_{N-1})$ are treated as unknown parameters that are estimated from our term structure data and that they imply a time-invariant one-factor model. The model is expressed in state-space form, which makes it possible to apply Kalman Filter techniques. Furthermore, it allows for the simultaneous inclusion of time series effects and cross sectional effects of the risk premium. The excess holding period return is divided into a component ζ_{t+1}, which only depends on time, and a component $Z(\tau)$, which only depends on maturity. The process $\{E_t\zeta_{t+1}\}_{t=1}^{T}$ shows evolution of the risk premium over time, whereas $\{Z(\tau)\}_{\tau=\tau_1}^{\tau_N}$ shows the relation between risk premia of bonds with different time to maturity, where the long term yield serves as the reference point.

Altogether our fixed maturity/random time effects panel model is given by

$$\begin{aligned} Y_{t+1} &= F_{t,t+1}\beta - Z\lambda_{t+1} + \eta_{t+1} \\ \lambda_{t+1} &= (1-\rho)\mu + \rho\lambda_t + \nu_{t+1} \\ \eta_{t+1} &\sim N(0, \Sigma) \\ \nu_{t+1} &\sim N(0, \sigma^2) \end{aligned} \tag{3.22}$$

where ρ, μ and σ^2 are the parameters in the AR(1) process and Σ denotes the covariance matrix of the cross sectional error term. We assume that η_{t+1} and ν_{t+1} are independent. Once the process for the excess holding period return has been estimated, the process for the risk premium at different maturities follows directly from the relations stated in equations (3.5) and in (3.21).

Of course, the error terms η_{t+1} are not cross-sectionally independent. We model the cross-sectional error term following the specification in chapter four. In order to keep the number of parameters to be estimated tractable, and to gain further efficiency, we assume that the correlation between error terms $\eta_{t+1}(\tau_i)$ and $\eta_{t+1}(\tau_j)$ depends on the distance

between the terms to maturity:

$$\mathsf{Corr}\left(\eta_{t+1}(\tau_i), \eta_{t+1}(\tau_j)\right) = \phi^{|\tau_i - \tau_j|} \tag{3.23}$$

with $0 < \varphi < 1$. Yields that are very close show high correlation, whereas yields that are far apart are less correlated. The specification resembles a cross-sectional AR(1) error term.

Besides cross sectional correlation we also account for possible cross sectional heteroskedasticity. Long term yields show less variance than short term yields. For that reason we specify the variance of the error term, $\eta_{t+1}(\tau)$, as a function of τ

$$\mathsf{Var}\left(\eta_{t+1}(\tau)\right) = \omega^2 \tau^{-2d}\,, \tag{3.24}$$

where ω^2 is a scale parameter, and d determines the sensitivity of the variance for the term to maturity. We estimate d along with the other parameters. If $d = 0$ then the error terms are homoskedastic in a model for the yields. In case $d = 1$, the model is homoskedastic in a model for (log-) bond prices. The heteroskedasticity implies a weighting scheme on the maturities. With $d > 0$ more emphasis is put on long term yields.

Altogether the cross sectional covariance matrix of the error terms for a model with maturities $\tau = \tau_1, \ldots, \tau_N$ is parametrized by the three parameters ω, ϕ and d. In matrix form the covariance structure for the cross sectional error terms is

$$\Sigma = \omega^2 S(\phi, d) \tag{3.25}$$

and

$$S(\phi, d) = \begin{bmatrix} \tau_1^{-2d} & \frac{\phi^{|\tau_2 - \tau_1|}}{(\tau_2 \tau_1)^d} & \cdots & \frac{\phi^{|\tau_N - \tau_1|}}{(\tau_N \tau_1)^d} \\ \frac{\phi^{|\tau_2 - \tau_1|}}{(\tau_1 \tau_2)^d} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{\phi^{|\tau_N - \tau_{N-1}|}}{(\tau_N \tau_{N-1})^d} \\ \frac{\phi^{|\tau_N - \tau_1|}}{(\tau_1 \tau_N)^d} & \cdots & \frac{\phi^{|\tau_N - \tau_{N-1}|}}{(\tau_{N-1} \tau_N)^d} & \tau_N^{-2d} \end{bmatrix} \tag{3.26}$$

where Σ and S are matrices of order $(N \times N)$. We assume that the same covariance matrix is applicable in the regression equation for the pure expectations hypothesis of equation (3.8), where the error term is ε_{t+1}.

3.4 Data

Interest rates are only directly observable for short maturities, implied by money market instruments. The longer term yields are implicitly available in coupon bearing bond prices. Because the expectations hypothesis is stated in terms of discount yields and discount forward rates, the analysis is greatly simplified if we can work with discount rates.

The alternative is to use approximate linear relations as proposed by Campbell (1986). Especially in periods of high and volatile interest rates care must be taken in applying the linear approximation.

In the literature there exist different methods to construct discount rates from coupon bearing bonds. See chapter two for an analysis of the Nelson-Siegel and the cubic spline methods. Here we follow McCulloch (1975), who introduces cubic spline functions. A cubic spline is a functional form for the discount function like

$$\delta_t(\tau) = 1 + \alpha_1\tau + \alpha_2\tau^2 + \sum_{j=1}^{L} \alpha_{2+j}(\tau - c_j)_+^3 \tag{3.27}$$

where $\delta_t(\tau)$ denotes the discount function at time t for a cash flow of \$1 with maturity τ, c_j, $j = 1, \ldots, L$ are break-points, $(\cdot)_+ = \max[.,0]$ and α_j, $j = 1, \ldots, L+2$ denote the parameters to be estimated. The functional form for the discount function at time t is found by minimizing

$$\sum_{i=1}^{K} \left[P_{it} - \sum_{j=1}^{J_i} C_{ij}\delta_t(\tau_{ij}) \right]^2 \tag{3.28}$$

where K is the number of bonds, P_{it} denotes the price of bond i at time t, corrected for accrued interest, C_{ij} denotes the cash flow of bond i with term τ_{ij} and J_i is the number of cash flows for bond i. Finally, the yield curve at time t is obtained by

$$Y_t(\tau) = -\frac{\ln[\delta_t(\tau)]}{\tau} \tag{3.29}$$

The splines are an a-theoretical way to transform the information in bond prices to yield curve data. We have used US government bond data, available from the CRSP tapes for the period January 1970 until December 1994. These data are available on a monthly basis. For maturities less than a year we also include the observed Treasury bill rates. For each month the spline function contains 12 breakpoints, and thus 14 unrestricted parameters α_j. The parameters of the spline function are different for every month. As the fitted curve does not perfectly fit all available bond prices, the method already filters out some error terms implicit in the quadratic fit criterium (3.28). Using the estimated cubic spline functions, we have constructed a panel which includes time series of yields with maturities of 1 to 6 months and 1 to 10 years, a total of 16 time series. The number of maturities that we include in the sample corresponds with the number of parameters in the spline functions. In essence the data on a large cross section of bond prices are condensed to 16 observations. At every time t the yield curve is represented by a cross section of yields, denoted by $Y_t(\tau_1), \ldots, Y_t(\tau_N)$, where $N = 16$.

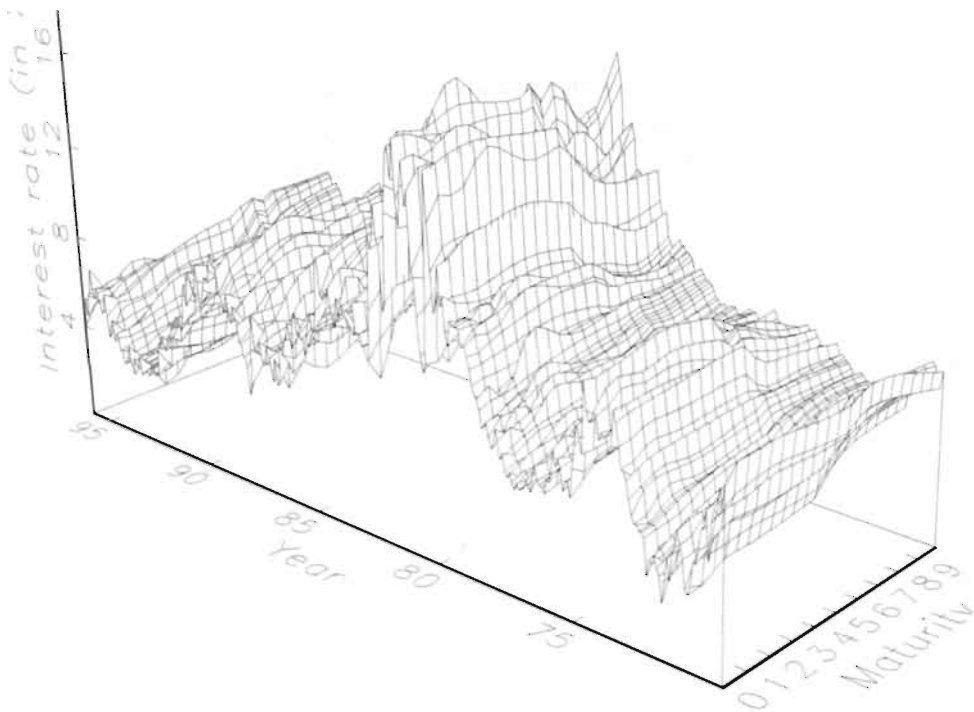

FIGURE 3.1: US INTEREST RATES

Figure 3.1 shows the full data panel. A number of issues are of interest. First, most of the yield curves are increasing. Second, in the period 1979-1984 the interest rates have been substantially larger than in the rest of the data sample. Third, for short maturities the yield curve is less smooth than for long maturities. Data at the short end of the yield curve are less reliable since they include monetary policy effects and the like.

In the top panel of table 3.1 summary statistics are presented for yields. The yield levels show an average term structure that is increasing. The term structure of volatilities is decreasing and from the last column we find that yields are highly autocorrelated. The bottom part of table 3.1 present the difference of yields and lagged forward rates. Under the pure expectations hypothesis the lagged forward rate is the expected value of the current yield. The second column denotes the average difference between the lagged forward rate and the current yield. It follows that next month's one month yield is on average 13 basis points lower than the current forward rate, next period's ten year yield is on average only 1 basis point lower than the current forward rate. A possible difference between the two is attributed to a risk premium. From equation (3.7) it follows that the average risk premium is given by

$$avg\left[\psi_t(\tau)\right] = \tau\left\{avg\left[F_{t,t+1}(\tau)\right] - avg\left[Y_{t+1}(\tau)\right]\right\} \tag{3.30}$$

In the last column the average risk premium is reported which follows from multiplication of the first and the second column. The statistics imply that the risk premium is on average increasing with maturity, the same holds for the associated standard errors. Equation (3.5) presents a definition of risk premium in a holding period return context. It follows that the average expected monthly excess return on a ten year bond over the risk free rate is 13 basis points.

TABLE 3.1: SUMMARY STATISTICS

(A) Yield Levels: $Y_t(\tau)$

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto Correlation
1 month	6.99	2.98	2.49	16.97	0.904
2 months	7.05	2.87	2.52	16.37	0.962
3 months	7.13	2.83	2.64	15.93	0.971
4 months	7.21	2.82	2.90	15.96	0.968
5 months	7.30	2.82	2.89	16.21	0.962
6 months	7.38	2.83	2.92	16.55	0.960
1 year	7.54	2.65	2.94	15.89	0.968
2 years	7.89	2.52	3.77	16.31	0.973
3 years	8.07	2.39	4.20	15.88	0.975
4 years	8.23	2.33	4.52	16.00	0.978
5 years	8.33	2.23	4.81	15.50	0.979
6 years	8.42	2.16	4.99	14.93	0.980
7 years	8.49	2.14	5.16	14.71	0.982
8 years	8.54	2.14	5.32	14.80	0.983
9 years	8.56	2.13	5.47	15.04	0.983
10 years	8.58	2.12	5.57	15.28	0.983

(B) Lagged Forward Rates minus Yields : $F_{t-1,t}(\tau) - Y_t(\tau)$

Maturity	Average	Standard deviation	Minimum	Maximum	Auto Correlation	Average Risk Premium
1 month	0.13	1.11	-3.13	4.61	0.300	0.01
2 months	0.14	0.88	-2.77	4.75	0.287	0.02
3 months	0.16	0.84	-2.69	4.85	0.091	0.04
4 months	0.17	0.85	-2.55	4.87	-0.014	0.06
5 months	0.16	0.87	-2.47	4.82	-0.070	0.07
6 months	0.13	0.85	-2.72	4.72	-0.053	0.07
1 year	0.09	0.70	-2.92	4.21	0.113	0.09
2 years	0.05	0.59	-2.85	3.53	0.161	0.10
3 years	0.05	0.54	-2.42	2.78	0.115	0.15
4 years	0.04	0.50	-2.18	2.29	0.152	0.14
5 years	0.03	0.46	-2.08	1.93	0.160	0.16
6 years	0.03	0.43	-1.95	1.68	0.124	0.16
7 years	0.02	0.41	-1.78	1.59	0.083	0.16
8 years	0.02	0.40	-1.58	1.58	0.066	0.15
9 years	0.02	0.39	-1.38	1.57	0.079	0.14
10 years	0.01	0.39	-1.17	1.57	0.089	0.13

TABLE 3.2: CORRELATION MATRIX

(A) Yield Levels: $Y_t(\tau)$

τ	1/12	2/12	3/12	4/12	5/12	6/12	1	2	3	4	5	6	7	8	9	10
1/12	1.00	0.98	0.96	0.95	0.95	0.95	0.94	0.91	0.90	0.88	0.87	0.85	0.83	0.82	0.82	0.82
2/12	0.98	1.00	0.99	0.99	0.98	0.98	0.97	0.95	0.93	0.91	0.89	0.87	0.86	0.84	0.84	0.85
3/12	0.96	0.99	1.00	1.00	0.99	0.99	0.98	0.96	0.93	0.92	0.90	0.88	0.86	0.85	0.85	0.85
4/12	0.95	0.99	1.00	1.00	1.00	1.00	0.99	0.96	0.94	0.92	0.91	0.89	0.87	0.86	0.85	0.86
5/12	0.95	0.98	0.99	1.00	1.00	1.00	0.99	0.97	0.94	0.93	0.91	0.89	0.87	0.86	0.86	0.86
6/12	0.95	0.98	0.99	1.00	1.00	1.00	0.99	0.97	0.95	0.93	0.91	0.89	0.88	0.87	0.86	0.87
1	0.94	0.97	0.98	0.99	0.99	0.99	1.00	0.99	0.97	0.96	0.95	0.93	0.91	0.90	0.90	0.90
2	0.91	0.95	0.96	0.96	0.97	0.97	0.99	1.00	0.99	0.99	0.98	0.97	0.96	0.95	0.95	0.95
3	0.90	0.93	0.93	0.94	0.94	0.95	0.97	0.99	1.00	1.00	0.99	0.99	0.98	0.97	0.97	0.97
4	0.88	0.91	0.92	0.92	0.93	0.93	0.96	0.99	1.00	1.00	1.00	0.99	0.99	0.98	0.98	0.98
5	0.87	0.89	0.90	0.91	0.91	0.91	0.95	0.98	0.99	1.00	1.00	1.00	0.99	0.99	0.99	0.99
6	0.85	0.87	0.88	0.89	0.89	0.89	0.93	0.97	0.99	0.99	1.00	1.00	1.00	1.00	0.99	0.99
7	0.83	0.86	0.86	0.87	0.87	0.88	0.91	0.96	0.98	0.99	0.99	1.00	1.00	1.00	1.00	0.99
8	0.82	0.84	0.85	0.86	0.86	0.87	0.90	0.95	0.97	0.98	0.99	1.00	1.00	1.00	1.00	0.99
9	0.82	0.84	0.85	0.85	0.86	0.86	0.90	0.95	0.97	0.98	0.99	0.99	1.00	1.00	1.00	1.00
10	0.82	0.85	0.85	0.86	0.87	0.90	0.95	0.97	0.98	0.99	0.99	0.99	1.00	1.00	1.00	1.00

(B) Lagged Forward Rates minus Yield: $F_{t-1,t}(\tau) - Y_t(\tau)$

τ	1/12	2/12	3/12	4/12	5/12	6/12	1	2	3	4	5	6	7	8	9	10
1/12	1.00	0.86	0.58	0.43	0.42	0.44	0.46	0.42	0.40	0.39	0.40	0.38	0.35	0.33	0.33	0.33
2/12	0.86	1.00	0.90	0.75	0.68	0.67	0.69	0.63	0.58	0.56	0.54	0.52	0.50	0.47	0.45	0.44
3/12	0.58	0.90	1.00	0.95	0.87	0.84	0.81	0.75	0.69	0.65	0.60	0.58	0.57	0.55	0.52	0.50
4/12	0.43	0.75	0.95	1.00	0.97	0.95	0.84	0.79	0.73	0.68	0.61	0.59	0.59	0.57	0.55	0.52
5/12	0.42	0.68	0.87	0.97	1.00	0.99	0.82	0.79	0.73	0.68	0.61	0.59	0.59	0.58	0.56	0.54
6/12	0.44	0.67	0.84	0.95	0.99	1.00	0.84	0.81	0.76	0.71	0.64	0.63	0.63	0.61	0.59	0.57
1	0.46	0.69	0.81	0.84	0.82	0.84	1.00	0.94	0.89	0.85	0.81	0.79	0.78	0.77	0.74	0.71
2	0.42	0.63	0.75	0.79	0.79	0.81	0.94	1.00	0.97	0.94	0.89	0.88	0.87	0.85	0.82	0.80
3	0.40	0.58	0.69	0.73	0.73	0.76	0.89	0.97	1.00	0.98	0.93	0.92	0.91	0.90	0.88	0.86
4	0.39	0.56	0.65	0.68	0.68	0.71	0.85	0.94	0.98	1.00	0.98	0.96	0.95	0.93	0.91	0.89
5	0.40	0.54	0.60	0.61	0.61	0.64	0.81	0.89	0.93	0.98	1.00	0.99	0.96	0.93	0.91	0.88
6	0.38	0.52	0.58	0.59	0.59	0.63	0.79	0.88	0.92	0.96	0.99	1.00	0.99	0.96	0.92	0.89
7	0.35	0.50	0.57	0.59	0.59	0.63	0.78	0.87	0.91	0.95	0.96	0.99	1.00	0.99	0.96	0.92
8	0.33	0.47	0.55	0.57	0.58	0.61	0.77	0.85	0.90	0.93	0.93	0.96	0.99	1.00	0.99	0.95
9	0.33	0.45	0.52	0.55	0.56	0.59	0.74	0.82	0.88	0.91	0.91	0.92	0.96	0.99	1.00	0.99
10	0.33	0.44	0.50	0.52	0.54	0.57	0.71	0.80	0.86	0.89	0.92	0.95	0.99	1.00	1.00	1.00

In table 3.2 another feature of yield curve data is illustrated. The top panel shows that yields that have maturities that are close together, are highly correlated. The correlation decreases as the difference between the maturities of the bonds increases. For the yield minus the lagged forward rate, we observe the same pattern although the correlations are not as high as in the case of yield levels. This feature of the data is important since it shows that the information available in the data is limited. We take this explicitly into account when we model the cross sectional relation between the error terms.

3.5 Estimation results

This section presents empirical results for the various models that deal with tests of the expectations hypothesis. We start with simple univariate regressions. Next, we pool the data and test the expectations hypothesis for all yields simultaneously. Then models are considered in which the risk premium is explicitly modeled. First we consider a fixed effects estimator, then we deal with the case where the risk premium is modeled as a combined fixed maturity/random time effects model.

Table 3.3 shows the univariate regression results of yields on lagged forward rates for all maturities, as given by equation (3.8). Typically, the slope coefficient β is close to one, which implies that the lagged forward rate seems a reasonable predictor of the future yield. For most maturities at the short end of the yield curve, the pure expectations hypothesis (*i.e.* $H_0 : \alpha = 0, \beta = 1$) is rejected, which follows from the LR statistic, that is compared with a critical value of $\chi^2_{0.95}(2) = 5.99$. It seems that the results for very short maturities up to 4 months are imprecise, probably caused by transitory noise at the short end, or because of liquidity problems around maturity dates of bonds. Bekaert, Hodrick and Marshall (1997) point out that rejection of the expectations hypothesis may very well be caused by small sample biases.

The Durbin-Watson statistic indicates that for some maturities there is residual autocorrelation. This may point at dynamics in the error term that were not modeled adequately, perhaps a time-varying risk premium. The residuals for the separate maturities are clearly heteroskedastic, given the increment in volatility of the residual term when the maturity of the yield is higher. We have also calculated the cross-sectional correlation between error terms (not reported) and find high correlation for residuals which differ only a little in time-to-maturity. The correlation decreases as the residuals differ more in time-to-maturity.

Remember that we employ panel data techniques in order to deal with two issues. First, pooling the data leads to more efficient estimates and will help to deal with the

TABLE 3.3: RESULTS UNIVARIATE REGRESSIONS

τ	α	$s.e.(\alpha)$	β	$s.e.(\beta)$	$\sigma[\epsilon(\tau))]$	LR	R^2	DW
$\frac{1}{12}$	0.039	0.173	0.976	0.023	0.092	1.10	0.863	1.39
$\frac{2}{12}$	0.136	0.139	0.962	0.018	0.146	4.60	0.907	1.38
$\frac{3}{12}$	0.159	0.132	0.957	0.017	0.208	6.46	0.914	1.75
$\frac{4}{12}$	0.194	0.136	0.950	0.017	0.280	8.31	0.911	1.96
$\frac{5}{12}$	0.211	0.140	0.950	0.018	0.358	8.16	0.908	2.07
$\frac{6}{12}$	0.175	0.139	0.959	0.017	0.420	5.67	0.912	2.05
1	0.090	0.124	0.977	0.015	0.692	2.27	0.932	1.74
2	0.086	0.114	0.983	0.014	1.174	1.55	0.946	1.65
3	0.117	0.110	0.979	0.013	1.602	2.50	0.950	1.74
4	0.106	0.107	0.982	0.012	1.984	1.90	0.955	1.67
5	0.010	0.105	0.984	0.012	2.315	1.68	0.957	1.65
6	0.110	0.101	0.983	0.012	2.593	1.96	0.960	1.72
7	0.117	0.098	0.983	0.011	2.870	2.22	0.963	1.81
8	0.116	0.095	0.984	0.010	3.176	2.14	0.966	1.84
9	0.113	0.094	0.985	0.011	3.506	2.03	0.967	1.82
10	0.113	0.094	0.985	0.011	3.855	1.91	0.967	1.80

Notes: The table shows the estimation results for the univariate regression of a yield with maturity τ, $Y_{t+1}(\tau)$ on a constant plus the forward rate $F_{t,t+1}(\tau)$. The constant is denoted as α, the slope is β. With $\sigma[\epsilon(\tau))]$ we denote the standard deviation of the error term. We present the corresponding standard errors, R^2 and Durbin-Watson statistic (DW). LR denotes the values of the likelihood ratio statistic for the joint test of $\alpha = 0$ and $\beta = 1$, which has a critical value of $\chi^2_{0.95}(2) = 5.99$. The sample consists of monthly observations for the period January 1970-December 1994.

small sample bias problem. Furthermore, panel data techniques allow us to model the risk premium explicitly, whereas in the univariate regression tests it is included in the error term. Now, we pool the data and rerun the regression model of equation (3.8), estimating the slope parameter for all maturities simultaneously. Also, we take into account the heteroskedasticity and the cross-sectional correlation by specifying the error term as proposed by equations (3.25) and (3.26). The results are presented in the first column of table 3.4. The slope parameter β is very close to one. Pooling the data also results in an even lower associated standard error, which still leads to a rejection of the expectations hypothesis. The covariance matrix of the error terms properly accounts for the heteroskedasticity in the residuals, see figure 3.2 for the term structure of residual volatilities, both as observed in the residuals and implied by the cross-sectional covariance matrix specification. Also the covariance matrix appears to model the correlation between residuals adequately. Residuals which differ only one month in maturity have a correlation coefficient of $\phi^{\frac{1}{12}} = 0.997$. Clearly, since ϕ is close to one, the correlation between residuals decreases smoothly with

difference in time-to-maturity, which is consistent with what we found in the univariate regression results.

TABLE 3.4: RESULTS POOLED REGRESSION

	Pooled	Maturity Effects
β	0.994	0.945
	(0.001)	(0.056)
ω	1.34	1.34
	(0.052)	(0.056)
ϕ	0.962	0.963
	(0.003)	(0.003)
d	-0.312	-0.323
	(0.003)	(0.003)
lnL	3306	3385

Notes: The table shows the estimation results for the pooled regression of $\tau Y_{t+1}(\tau)$ on $\tau F_{t,t+1}(\tau)$. With β we denote the regression coefficient. The covariance matrix is specified by the parameters ω, ϕ and d. lnL denotes the value of the loglikelihood function in the optimum. In the Pooled model, we pool the data for all yields with different maturities. The Maturity effects model incorporates a dummy parameter for each maturity. The sample consists of monthly observations for the period January 1970- December 1994. Standard errors are within parentheses.

In the second column of table 3.4 the parameter results for the fixed maturity effects model as specified by equation (3.16) are given. We find a drop in the estimate for β, which decreases to 0.945. The expectations hypothesis cannot be rejected in the case of fixed maturity effects. The standard errors are high enough, even after adjustment for low autocorrelation in the residuals, to favor the expectations hypothesis. The model explicitly includes the parameter $\psi_t(\tau)$ through inclusion of a dummy for each maturity. However, only if β is equal to 1, we can interpret the parameter $\psi_t(\tau)$ as a risk premium. Otherwise, we have to correct for the deviation of β from 1. Estimates for the parameter $\psi_t(\tau)$, along with the associated standard errors are given in table 3.5. In the last column we also quantify the risk premium, by taking account of the deviation. The levels of the risk premia are consistent with the historical average values we reported in the data section. In the case β is constrained to be 1, the fixed maturity estimator for the risk premium is nothing but the historical average value of yields minus lagged forwards, as given in the bottom part of table 3.1. The maturity effects show that risk premia increase with time-to-maturity τ. Testing whether the maturity effects are zero amounts to a likelihood

ratio (LR) test between the pooled model and the fixed maturities effects model. The hypothesis is clearly rejected, since the regression results report an increase of two times the loglikelihood of about 158, while the critical value equals $\chi^2_{0.95}(16) = 26.30$. The LR test favors the presence of risk premia that vary with time-to-maturity. Note that estimation results for the covariance specification of the error term are almost identical to the *pooled* case.

TABLE 3.5: RESULTS FIXED MATURITY EFFECTS: $\psi(\tau)$

τ	$\psi(\tau)$	se($\psi(\tau)$)	Risk Premium
$\frac{1}{12}$	0.022	0.035	0.01
$\frac{2}{12}$	0.042	0.044	0.02
$\frac{3}{12}$	0.061	0.051	0.04
$\frac{4}{12}$	0.078	0.056	0.06
$\frac{5}{12}$	0.103	0.061	0.07
$\frac{6}{12}$	0.139	0.065	0.07
1	0.334	0.087	0.09
2	0.775	0.126	0.10
3	1.190	0.166	0.15
4	1.675	0.207	0.14
5	2.142	0.249	0.16
6	2.624	0.292	0.16
7	3.117	0.335	0.16
8	3.613	0.378	0.15
9	4.104	0.422	0.14
10	4.596	0.465	0.13

Notes: The table shows the estimation results for the fixed maturity effects $\psi(\tau)$ and the associated standard errors in the case of the pooled regressions with fixed individual effects. The third column quantifies the risk premium. The sample consists of monthly observations for the period January 1970 - December 1994.

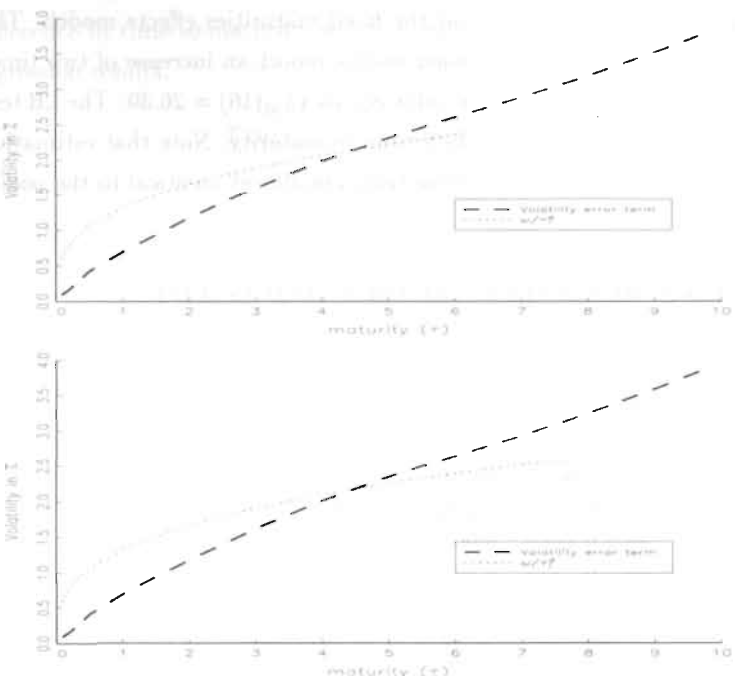

FIGURE 3.2: VOLATILITIES ERROR TERM

Each of the panels shows the implied volatility functions of the error terms for the panel data models. The top panel shows results for the pooled regression model and the bottom figure for the fixed maturity effects model.

In table 3.6 we report the estimation results for the full model that was presented in equations (3.22), (3.23) and (3.24). The model incorporates both fixed maturity effects and random time effects. The model is written in state space form and is estimated using the time-invariant Kalman filter. Model I in the first column of table 3.6 shows the regression results for the most general case. The slope parameter, β, shows a further decrease and is far away from the value where the expectations hypothesis would hold. The parameters ρ, μ and σ model the time series dynamics for ψ_t, which is the time series related to the longest maturity $\tau_N = 10$. We find moderate persistence in the process for ψ_t, with an AR(1) parameter of $\rho = 0.675$. Because β differs from 1, ψ_t cannot be interpreted as the risk premium for the long maturity yield. For the same reason, μ is not the average value of the long term risk premium. At time t, the bias is of the order $(1 - \beta)\,\tau_N F_{t,t+1}(\tau_N)$, which is on average 17.96. The risk premium time series processes for the other maturities are linked to ψ_t through scaling with the maturity parameter $Z(\tau)$, which is found in table 3.7. As in the fixed maturity effects case, we find that Z increases with maturity. In figure 3.3 both the time series pattern and the cross sectional relation are depicted. The covariance

matrix specification of the error term takes account of the high correlation between error terms that differ little in time to maturity. For error terms that differ 1 year in time to maturity, the implied correlation is $\hat{\phi} = 0.834$, for error terms that differ only one month in time to maturity the correlation is $\hat{\phi}^{\frac{1}{12}} = 0.985$, which corresponds to the data. The covariance specification is similar for restricted versions of the model that we consider next.

TABLE 3.6: RESULTS FIXED MATURITY/RANDOM TIME EFFECTS PANEL MODEL

	I	II
β	0.791	"1"
	(0.005)	-
ω	0.540	0.516
	(0.023)	(0.026)
ϕ	0.834	0.806
	(0.01)	(0.02)
d	-0.252	-0.246
	(0.004)	(0.005)
ρ	0.675	0.115
	(0.045)	(0.055)
μ	17.57	0.206
	(0.99)	(0.31)
σ	4.18	3.68
	(0.165)	(0.173)
lnL	4076	3960

Notes: The tabel reports estimation results for the pooled regression model with a covariance structure that takes account of the correlation between yields. An AR(1) process is incorporated for the 10 year yield risk premium. The remaining risk premia are related to this proces by multiplying with $Z(\tau)$. Model I is the general case, in model II we restrict β to 1. Standard errors are within parentheses.

Model II in table 3.6 deals with a restricted version of model I. Since under the expectation hypothesis the theoretical value of the slope coefficient is $\beta = 1$, we restrict the parameter to this value. We find a lower AR(1) parameter estimate in the random effects specification of the risk premium, but it is significantly different from zero. This specification suggests that the associated risk premium still increases with maturity, and that there is also small but significant predictive power with respect to the time series dimension of the risk premium. This average value for the risk premium, $\mu = 0.206$, is similar to what we found in the fixed maturity case. The likelihood ratio test favors a model with time-varying risk premia and, thus, rejection of the expectations hypothesis, since the LR statistic is 232, whereas the critical value is $\chi^2_{0.95}(1) = 3.84$.

TABLE 3.7: RESULTS RISK PREMIUM SPECIFICATION: $Z(\tau)$

τ	I (pars)	I (s.e.)	II (pars)	II (s.e.)
$\frac{1}{12}$	0.007	(0.008)	0.008	(0.044)
$\frac{2}{12}$	0.013	(0.008)	0.019	(0.043)
$\frac{3}{12}$	0.020	(0.011)	0.033	(0.062)
$\frac{4}{12}$	0.027	(0.009)	0.047	(0.051)
$\frac{5}{12}$	0.034	(0.011)	0.061	(0.060)
$\frac{6}{12}$	0.042	(0.011)	0.074	(0.059)
1	0.088	(0.012)	0.148	(0.061)
2	0.186	(0.011)	0.275	(0.057)
3	0.284	(0.010)	0.395	(0.048)
4	0.387	(0.009)	0.499	(0.044)
5	0.487	(0.007)	0.575	(0.035)
6	0.589	(0.010)	0.649	(0.049)
7	0.692	(0.011)	0.732	(0.064)
8	0.795	(0.009)	0.823	(0.053)
9	0.898	(0.005)	0.917	(0.032)
10	1	-	1	-

Notes: The tabel reports the estimation results for $Z(\tau)$, which relates the risk premium of the yield to the risk premium process for the 10 year yield. Model I is the most general case and in model II we restrict β to 1. Standard errors are within parentheses.

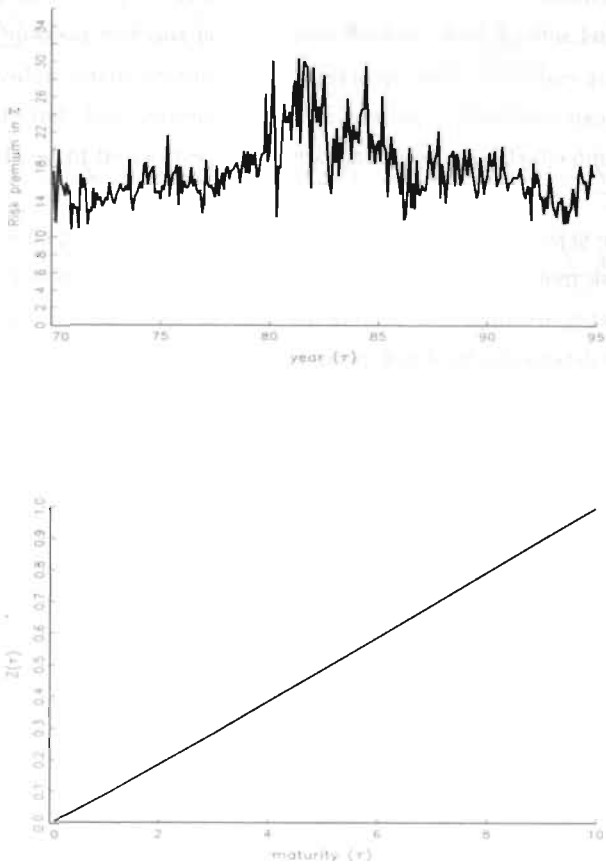

FIGURE 3.3: RISK PREMIUM SPECIFICATION

The top figure shows the time series process for the risk premium related to the 10 year yield. In the bottom figure the cross sectional relation to this risk premium is given.

3.6 Concluding Remarks

In this chapter we have developed a panel data model for the term structure of interest rates which combines fixed maturity effects with random time effects. Relative to standard regression approaches, an important advantage of the model is that it allows us to explicitly disentangle risk premia and unexpected excess returns. The panel setting also helps to

mitigate small sample bias and to increase statistical efficiency in testing procedures.

Our empirical results demonstrate a resounding rejection of the expectations hypothesis in the multivariate panel setting, even though the rejection was not possible for the univariate models at the long end of the term structure. The point estimates indicate that a considerable degree of mean reversion is present in the risk premia and that the slope coefficient β,which is one under the expectations hypothesis, is estimated to be about 0.8 for the full data panel.

The results quantify the seize of the risk premium. They also shed light on changing compensation above the risk free rate. We find that in periods of high interest rates and in periods when forward rates are high risk premia are high as well. Hence, the results provide new insights in the determinants of risk premia.

As a forecasting model of future yields the model that includes the time variation in the risk premia provides a more complete picture than the model that only incorporates the forward rate as an explanatory variable.

Appendix

In this appendix we derive the first order conditions which are necessary to obtain parameter estimates. Furthermore we derive the Hessian matrix which is used to compute standard errors of the parameter estimates. We consider the case where the risk premia are treated as individual effects. The loglikelihood reads

$$\ln L = -\frac{1}{2}\ln|S| - \frac{1}{2}NT\ln(\omega^2) - \frac{1}{2\omega^2}\sum_{t=1}^{T-1}(Y_{t+1} - F_{t,t+1}\beta - \psi)' S^{-1}(Y_{t+1} - F_{t,t+1}\beta - \psi)$$

The first derivative with respect to ψ is

$$\frac{\partial \ln L}{\partial \psi} = \frac{1}{\omega^2}\sum_{t=1}^{T-1}(Y_{t+1} - F_{t,t+1}\beta - \psi)' S^{-1}$$

The first derivative with respect to β is

$$\frac{\partial \ln L}{\partial \beta} = \frac{1}{\omega^2}\sum_{t=1}^{T-1}(Y_{t+1} - F_{t,t+1}\beta - \psi)' S^{-1} F_{t,t+1}$$

For $\hat{\psi}$ it therefore holds that

$$\hat{\psi} = \frac{1}{T-1}\sum_{t=1}^{T-1}(Y_{t+1} - F_{t,t+1}\beta)$$

or

$$\hat{\psi} = \left(\frac{1}{T-1}\sum_{t=1}^{T-1} Y_{t+1}\right) - \left(\frac{1}{T-1}\sum_{t=1}^{T-1} F_{t,t+1}\right)\hat{\beta}$$

For $\hat{\beta}$ it holds that

$$\hat{\beta} = \left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1} F_{t,t+1}\right)^{-1} \left(\sum_{t=1}^{T-1} Y'_{t+1} S^{-1} F_{t,t+1}\right) - \left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1} F_{t,t+1}\right)^{-1} \left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1}\right) \hat{\psi}$$

or

$$\begin{aligned}\hat{\beta} &= \left[I - \left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1} F_{t,t+1}\right)^{-1} \left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1}\right) \left(\frac{1}{T-1} \sum_{t=1}^{T-1} F_{t,t+1}\right)\right]^{-1} \left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1} F_{t,t+1}\right)^{-1} \\ &\quad \times \left[\left(\sum_{t=1}^{T-1} Y'_{t+1} S^{-1} F_{t,t+1}\right) - \left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1}\right) \left(\frac{1}{T-1} \sum_{t=1}^{T-1} Y_{t+1}\right)\right]\end{aligned}$$

$$\begin{aligned}\hat{\beta} &= \left[\left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1} F_{t,t+1}\right) - \left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1}\right) \left(\frac{1}{T-1} \sum_{t=1}^{T-1} F_{t,t+1}\right)\right]^{-1} \times \\ &\quad \times \left[\left(\sum_{t=1}^{T-1} Y'_{t+1} S^{-1} F_{t,t+1}\right) - \left(\sum_{t=1}^{T-1} F_{t,t+1} S^{-1}\right) \left(\frac{1}{T-1} \sum_{t=1}^{T-1} Y_{t+1}\right)\right]\end{aligned}$$

The second derivatives are required to calculate standard errors for the parameter estimates.

$$\frac{\partial^2 \ln L}{\partial \psi \partial \psi'} = -\frac{1}{\omega^2} (T-1) S^{-1}$$

$$\frac{\partial^2 \ln L}{\partial \beta^2} = -\frac{1}{\omega^2} \sum_{t=1}^{T-1} F_{t,t+1} S^{-1} F_{t,t+1}$$

$$\frac{\partial^2 \ln L}{\partial \psi \partial \beta} = -\frac{1}{\omega^2} \sum_{t-1}^{T-1} F_{t,t+1} S^{-1}$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \psi'} = -\frac{1}{\omega^2} \sum_{t-1}^{T-1} S^{-1} F_{t,t+1}$$

and the covariance matrix of the parameter estimates reads

$$cov(\beta, \psi) = \omega^2 \begin{bmatrix} \sum_{t=1}^{T-1} F_{t,t+1} S^{-1} F_{t,t+1} & \sum_{t=1}^{T-1} F_{t,t+1} S^{-1} \\ \sum_{t=1}^{T-1} S^{-1} F_{t,t+1} & (T-1) S^{-1} \end{bmatrix}^{-1}$$

Chapter 4

Direct Estimation of the Risk Neutral Factor Dynamics of Affine Term Structure Models

4.1 Introduction

Models of the term structure of interest rates provide the shape and dynamics of the yield curve from assumptions about the dynamics of some underlying factors and the price of the risk associated with each factor. A limited number of underlying state variables are assumed to account for the behaviour of interest rates of many different maturities. The models are often applied in pricing various interest rate derivative securities. These applications require an estimate of the risk adjusted (or risk neutral) dynamics of the factors.

The risk neutral process can be inferred from the observed term structure of interest rates, or from its change relative to an initial term structure. To estimate the parameters of the process one would like to have an analytically tractable model that relates the yields to the factors. A class of models that has received a lot of attention is the affine class, analyzed in detail by Duffie and Kan (1996), Campbell, Lo and MacKinlay (1997, ch 11) and Frachot and Lesne (1993). It has a tractable linear structure and nests a number of well-known equilibrium term structure models like the Vasicek (1977) and Cox, Ingersoll and Ross (1985, CIR) model, and their multifactor generalizations..

Even though the affine class is tractable, direct estimation of the risk neutral parameters has not been very successful.. Many studies have attempted to estimate parameters from cross sectional yield curve data. In these papers the structural parameters are re-estimated every time period using the bond prices at that moment. All of them report unstable and

erratic parameter estimates.[1] The econometric problems could be due either to a lack of information in yield curve data for a single date, or to misspecification of the models.

The aim of this paper is to revisit the direct estimation of the risk neutral process using panel data techniques. We pool monthly yield curves for a period of twenty-five years to reduce the problem of low power. As in the cross sectional studies cited above, the parameters are estimated without explicit reference to the time series properties of interest rates. Econometrically, this leads to a panel model with time dependent fixed effects.

From the cross sectional estimation we obtain the implied dynamic process of interest rates that generates yield curves that closely fit the observed term structure. But rather than modeling the levels of the yield curve, applications to derivative pricing emphasize models of the yield curve in deviation of last period's term structure.[2] These models provide an expression for the shape of the current term structure, conditional on an exogenously given initial term structure. This initial term structure need not, but could, be consistent with an equilibrium model. Some of these models, for example the extended Vasicek and extended CIR models, belong to the extended affine class as analyzed in Frachot and Lesne (1993). Our panel data analysis with fixed effects is suited for both types of term structure models. The equilibrium and extended models will generally emphasize different moments of the data. But depending on the properties of measurement error in the empirical model, the two formulations are shown to be closely related, and can be compared in a common econometric framework.

Following Jacquier and Jarrow (1998), Brown and Dybvig (1986) and others, we add an error term to the theoretical model. The covariance matrix of the error term is modeled explicitly to take account of heteroskedasticity and the cross sectional correlation of yields. We tightly parameterize the cross sectional covariance structure, since we wish to include a broad range of maturities in the empirical model. Moreover, the specification explicitly deals with residual autocorrelation.

We find that the efficiency gains of the panel are large. For the one- and two-factor models that we estimate we obtain parameter estimates with very low standard errors. The first factor has very slow mean reversion, but with a t-ratio of more than eight mean reversion is significantly different from zero. The average convexity of the yield curve provides a sharp point estimate of the implied volatility of the factors. We also find that the extended and equilibrium models give very similar point estimates for the parameters of interest.

[1] See for example Brown and Dybvig (1986), Brown and Schaefer (1994), Dahlquist and Svensson (1994), DeMunnik and Schotman (1994), Addolorato and Berardi (1994), and Sercu and Wu (1997).

[2] See Hull and White (1990) and Heath, Jarrow and Morton (1992) for a theoretical analysis. See Jarrow (1996) and Hull (1996) for a textbook treatment.

Our fixed effects model contrasts with other panel data studies of affine term structure models.[3] In these studies the cross sectional yield curve model is estimated jointly with a time series process for the factors. Combining cross section and time series information exploits even more information in the data. But time series data are related to the actual dynamics, so that additional assumptions on the factor risk prices must be made. The usual assumptions are that the price of risk is either constant, or proportional to (one of) the factors. For the fixed effects analysis we do not need any assumption on the actual time series properties of the factors. Only the risk neutral process matters. Still we also obtain sharp estimates of the parameters, in particular mean reversion. Empirically, 25 years of pooled cross sections dominate the time series information.

The panel data analysis also allows a detailed misspecification analysis. Misspecification here refers to the shape of the yield curve, in contrast to the dynamic structure of interest rate processes. When we find misspecification for the yield curve, this misspecification will be present a fortiori in the panel models that combine time series and cross sectional information. For two-factor Gaussian models we can not reject the cross sectional restrictions. The two factor model is very parsimonious, containing only five structural parameters that are sufficient to describe the variety of shapes encountered in a history of twenty five years of monthly term structures. In addition to the five structural parameters, each yield curve in a two factor model has two time dependent parameters, only relevant to that particular yield curve.

For those models that appear correctly specified in the cross sectional dimension, it is interesting to compare our results to the other panel literature by imposing assumptions on the price of risk and the time series behavior of the factors. In a broad sense, the cross sectional information is consistent with recent empirical time series studies. Andersen and Lund (1996) and Balduzzi, Das and Foresi (1998) argue that interest rate dynamics contains at least two factors. One of the factors is a slowly evolving mean, and the second factor represents quickly mean reverting movements around this mean.

The remainder of this chapter is organized as follows. Section 4.2 introduces notation and provides a brief review of affine term structure models. Section 4.3 considers the special case of the (extended) Vasicek model in detail. In section 4.4 we illustrate the problems of cross sectional estimation, and explain why pooling different cross sections will lead to a very efficient estimator. Section 4.5 discusses the econometric panel data model. The final part of this section discusses the misspecification tests. Section 4.6 describes the data and stylized facts of the US term structure. Section 4.7 reports the empirical results.

[3] See for example Chen and Scott (1993, 1995), Gibbons and Ramaswamy (1993), DeJong (1997), Geyer and Pichler (1997), Frühwirth-Schnatter and Geyer (1997), Babbs and Nowman (1998), Lund (1997), Kappi (1997), Duan and Simonato (1995), Gong and Remolona (1996).

Section 4.8 concludes.

4.2 Affine term structure models

Duffie and Kan (1996) consider equilibrium formulations of the affine yield curve models. Frachot and Lesne (1993) generalize this approach to deal with the extended specifications, that match an initial yield curve exactly.[4] The general idea underlying the affine term structure models is that at time t there exist a vector of K factors, Z_t, that govern the term structure movements. The drift and the diffusion process of the factors are affine in the factors. The affine class has gained popularity because it leads to a tractable solution for the term structure of interest rates that is itself an affine function of the factors.

The risk adjusted process for the underlying factors in an affine term structure model is specified as

$$dZ_t = (\Phi_t - \Gamma Z_t)\, dt + V(Z_t) dW_t \tag{4.1}$$

where Z_t denotes the vector of K factors, W_t is a K-dimensional Brownian motion under the risk-adjusted probability measure, and Φ_t is a $(K \times 1)$ deterministic function. The time variation in Φ_t enables an exact fit of an observed initial yield curve and hence represents the no-arbitrage case. When Φ_t is a vector of constants we are in the equilibrium model case of Duffie and Kan (1996). Mean reversion is determined by the $(K \times K)$ matrix Γ, and $V_t(Z_t)$ is defined such that the covariance matrix is affine in the factors

$$\Psi_t = V_t(Z_t) V_t(Z_t)' = \begin{pmatrix} \alpha_{11} Z_t + \beta_{11} & \cdots & \alpha_{K1} Z_t + \beta_{K1} \\ \vdots & & \vdots \\ \alpha_{K1} Z_t + \beta_{K1} & \cdots & \alpha_{KK} Z_t + \beta_{KK} \end{pmatrix} \tag{4.2}$$

where α_{ij} are $(1 \times K)$ vectors of parameters, and β_{ij} are scalars.

Examples of affine term structure models are the equilibrium single factor models of Vasicek (1977) and Cox, Ingersoll and Ross (1985). In these models the single factor Z_t is the instantaneous risk free rate, r_t. For $K > 1$, the models include the multifactor versions of the Vasicek and CIR models, and some generalizations.. Further examples are discussed in Chen and Scott (1993) and Duffie and Kan (1996). Examples of the no-arbitrage class are the extended models of Vasicek and CIR considered by Hull and White (1990). Campbell, Lo and MacKinlay (1997) develop the affine models in discrete time.

[4] Throughout the chapter we will use the term "equilibrium" model for those models that endogenously determine the term structure by no-arbitrage arguments. Strictly speaking, not all affine models with constant parameters have a general equilibrium justification. Likewise, we will use the term "extended" for all models that start with an exogenously given yield curve. Sometimes these models are also called "evolutionary".

Define $Y_t(\tau)$ as the yield of a discount bond at time t with maturity τ. Frachot and Lesne (1993) show that with this specification of the factor dynamics no-arbitrage arguments imply a yield curve that is affine in the factors,

$$Y_t(\tau) = A_t(\tau) + B(\tau)Z_t \tag{4.3}$$

where $A_t(\tau)$ is a scalar and $B(\tau)$ a $(1 \times K)$ vector; both are functions of the structural parameters $\pi = (\alpha, \beta, \Phi_t, \Gamma)$ in the factor specification. In general the functions $A_t(\tau)$ and $B(\tau)$ are found by numerically solving a set of Ricatti differential (or difference in discrete time) equations. Next section considers the explicit functional relation for the special case of Gaussian factor models.

For pricing of interest rate derivatives the function $B(\tau)$ is the primary object of interest in term structure models. As shown in Heath, Jarrow and Morton (1992) and also explained in Hull (1996) all interest rate derivative securities can be priced with as inputs the "volatility" function $B(\tau)$, the factor covariance matrix Ψ_t, and the initial term structure. The purpose of the econometric analysis of yield curve models is the estimation of the function $B(\tau)$. For this the cross sectional model (4.3) suffices, and all we need is an assumption about the risk neutral factor dynamics.

The covariance matrix Ψ_t can often be estimated more precisely from time series data. Time series descriptions of the yield curve would focus on the factor process in (4.1). The general representation of the associated process under the actual probability measure reads

$$dZ_t = (\Phi_t - \lambda_t - \Gamma Z_t)\,dt + V(Z_t)d\tilde{W}_t \tag{4.4}$$

where λ_t is a vector of K prices of risk. Using time series data for statistical inference entails additional assumptions about the price of risk. In an equilibrium framework the risk prices are determined by preferences of agents, or more generally through a pricing kernel. While not every time variation in λ_t is allowed, the theory does not specify how they vary.

Before we develop an econometric model, it is useful to discuss the calibration practice of the extended models. The functions Φ_t and $A_t(\tau)$ have a time t subscript, but both are deterministic functions of time, whose time variation is fully determined by the observed term structure $Y_{t_0}(\tau)$ at date t_0, when the model was calibrated.

Over a short period the yield curve at time $t_0 + \Delta t$ will be close to the calibrated yield curve at time t_0 due to the smooth nature of the Brownian motion process. Over longer horizons the yield curve can wander away further from the initial yield curve. The variance of the distance is controlled by the integrated factor dynamics from t_0 to the current time t. Since the factor process is mean reverting the effect of the initial conditions will fade out

gradually. Mean reversion implies that long term yields and forward rates both converge to a constant, and this in turn implies that, when t_0 is fixed, $\lim_{t\to\infty} \Phi_t = \bar{\Phi}$ is a constant, and $A_t(\tau)$ will converge to a function of τ only. If the limit of Φ_t from the calibrated yield curve is the same constant as in the equilibrium formulation ($\bar{\Phi} = \Phi$), then the yield curves generated by the extended models will converge to the yield curves from equilibrium formulations.

For the econometric inference we will use observed yield curves over a long period of time. If the extended model is calibrated only once, say in January 1960, then the yield curves in the nineties will only be affected by the initial calibration by the level parameter $\bar{\Phi}$. In that case the extended model is in fact nothing but a restricted version of the equilibrium models. In the equilibrium models Φ is a free parameter, whereas it would be exogenously specified by more than thirty year old data in the extended model.

However, common practice in applications of the no-arbitrage models is repeated calibration, instead of a one time calibration. The model is calibrated every time we get a new term structure. With monthly data the functions Φ_t and $A_t(\tau)$ are updated every month, so that $t_0 = t - 1$ and $A_t(\tau)$ becomes a function of data from last month's term structure.

If, for the purpose of parameter estimation, the model is calibrated repeatedly, then the equilibrium and extended versions of the model only coincide if last month's term structure happens to be consistent with an equilibrium formulation. The extended model therefore allows for deviations from the unconditional equilibrium model that are expected to disappear gradually in a way that is consistent with the no-arbitrage condition. Because the intercept $A_t(\tau)$ depends on data from time $t - 1$, parameters are estimated from movements of the yield curve relative to its previous shape. In that sense it is a more flexible model for empirical work than the equilibrium model. When we develop the econometric model we will always consider the extended models being calibrated every month. In contrast to the equilibrium model, where the parameters are estimated from the shape of the yield curve, the extended models look at the cross section of changes in the term structure.

Using results in El Karoui, Frachot and Geman (1998) and Frachot and Lesne (1993) the affine model can be written in the representation

$$Y_t(\tau) = F_{t_0,t}(\tau) + a(t - t_0, \tau) + B(\tau)z_t \tag{4.5}$$

where $F_{t_0,t}(\tau)$ is the forward rate at time t_0 for the contract period t to $t+\tau$, $a(t-t_0, \tau)$ is a function depending on the structural parameters, and z_t is a K-vector of factors (different from Z_t). Since for the econometric model we will assume $t_0 = t - 1$, the first argument of

$a(t - t_0, \tau)$ does not depend on time, so that

$$Y_t(\tau) = F_{t-1,t}(\tau) + a(\tau) + B(\tau)z_t \tag{4.6}$$

which is a representation with constant parameters. The extended model has the same factor loading $B(\tau)$ as the corresponding equilibrium model. But in the extended model the current yield curve $Y_t(\tau)$ is modelled conditional on the forward rate curve $F_{t-1,t}(\tau)$. The model has similar structure as the equilibrium model, but with yields in deviation of their corresponding forward rates instead of yield levels.

4.3 Gaussian models

In general the transformation from factor dynamics to the affine yield curve parameters $A_t(\tau)$ and $B(\tau)$ is not available in closed form. The linear Gaussian case with uncorrelated factors is an example where closed form solutions are available, where comparison between equilibrium and extended models is straightforward, and where the structural parameters each have a clearly distinguishable function in the model. While the econometric model will be developed for the general affine class, the empirical analysis will concentrate on the Gaussian case with uncorrelated factors.

The single factor equilibrium Gaussian case is the Vasicek (1977) model. In this case $K = 1$ and $\alpha_{11} = 0$ in (4.1). The instantaneous spot rate r is the factor driving the yield curve, with risk adjusted diffusion process of the form

$$dr_t = \kappa(\mu - r_t)dt + \sigma dW_t \tag{4.7}$$

where κ is the parameter of mean reversion, μ is the unconditional mean of r, and σ is the instantaneous standard deviation. A multiple factor generalization is given by

$$r_t = \sum_{j=1}^{K} z_{jt} \tag{4.8}$$

$$dz_{jt} = \kappa_j(\mu_j - z_{jt})dt + \sigma_j dW_{jt} \tag{4.9}$$

where the Brownian motions W_{jt} are mutually uncorrelated. The associated yield curve model is represented by

$$Y_t(\tau) = A(\tau) + \sum_{j=1}^{K} b_j(\tau)Z_{jt} \tag{4.10}$$

where

$$A(\tau) = \sum_{j=1}^{K} \left((1 - b_j(\tau))\theta_j + \frac{(1 - e^{-\kappa_j \tau})^2}{4\kappa_j^3 \tau}\sigma_j^2 \right)$$

$$b_j(\tau) = \frac{1 - \exp(-\kappa_j \tau)}{\kappa_j \tau}$$

where $\theta_j = \mu_j - \frac{\sigma_j^2}{2\kappa_j^2}$. The shape of $B(\tau) = (b_1'(\tau), \ldots, b_K'(\tau))$ depends solely on the mean reversion parameters κ_j. Each element $b_j(\tau)$ is monotonically decreasing with $b_j(0) = 1$ and $b_j(\infty) = 0$. As in all affine models, the long term discount yield $Y(\infty)$ is a constant. At every date the yield curve must converge to the same constant, irrespective of the initial location and shape. Note that in the multifactor model not all parameters are identified. Reparameterizing by $Z_{jt} = z_{jt} - \theta_j$, one can verify that we can only identify $\theta = \sum_{j=1}^K \theta_j = \lim_{\tau \to \infty} Y_t(\tau)$, but not the individual θ_j.[5] Also note that the model is not only affine in the factors Z_{jt}, but also in the parameters θ and σ_j^2. For cross sectional estimation, the only nonlinearity is in κ_j.

The parameters θ and σ_j^2 only show up in $A(\tau)$, so that $B(\tau)$ solely depends on the mean reversion. Since the factor loadings (or volatility structure) $B(\tau)$ is the object of interest for the purpose of derivative pricing, our main concern will be the estimation of the mean reversion parameters. This is a specific feature of the Gaussian models. In the general affine model, and also in the CIR model, the factor loadings depend on all structural parameters.

The extended Vasicek is similar to (4.7), but now μ is replaced by μ_t. The parameters μ_t fluctuate in such a way that they are consistent with the observed term structure characteristics at some initial date t_0. Calibrating the term structure at some initial time t_0, Hull and White (1993) derive expressions for the coefficients μ_t and $A_t(\tau)$ for $t > t_0$. Letting $t_0 = t - 1$, and defining h as the length of the interval between successive observations we use their results to obtain an expression analogous to (4.6),

$$Y_t(\tau) - F_{t-1,t}(\tau) = a(\tau) + B(\tau)(r_t - f_{t-1,t}) \tag{4.11}$$

where

$$a(\tau) = (1 - e^{-2\kappa h})\frac{(1 - e^{-\kappa\tau})^2}{4\kappa^3\tau}\sigma^2 \tag{4.12}$$

$F_{t-1,t}(\tau)$ is the forward rate defined below (4.5), and $f_{t-1,t}$ is the instantaneous forward rate at time $t-1$ relating to time t. The variables $Y_t(\tau) - F_{t-1,t}(\tau)$ and $r_t - f_{t-1,t}$ can be interpreted as unexpected shocks under the risk neutral probability measure. The main differences between the standard and the extended Vasicek models are the restrictions on the levels of the yields, which appear in the standard Vasicek model through the parameter θ, but not in the extended version.

From the results of Frachot and Lesne (1993) we obtain a multifactor generalization of (4.11) as

$$Y_t(\tau) - F_{t-1,t}(\tau) = a(\tau) + \sum_{j=1}^{K} b_j(\tau) Z_{jt} \tag{4.13}$$

[5] See Dai and Singleton (1997) for a detailed analysis of identification and alternative parameterizations.

where $b_j(\tau)$ is the same as in (4.10), and

$$a(\tau) = \sum_{j=1}^{K} (1 - e^{-2\kappa_j h}) \frac{(1 - e^{-\kappa_j \tau})^2}{4\kappa_j^3 \tau} \sigma_j^2$$

So far we only discussed the risk neutral factor process. To include time series properties we could make the additional assumption that the price of risk λ is a constant. For example, internal coherence of the one factor model then requires that the time series process of r_t must be similar to the spot rate process, given in (4.7),

$$dr = \tilde{\kappa}(\tilde{\mu} - r)dt + \tilde{\sigma} d\tilde{W}_t \tag{4.14}$$

The tildes on top of the structural parameters indicate that this process is specified under the original probability measure with Brownian motion $\tilde{W}$ instead of the risk neutral processes in (4.7). If the model is correctly specified the structural parameters differ from the risk neutral parameters only by the parameter $\lambda = \mu - \tilde{\mu}$, while $\tilde{\kappa} = \kappa$ and $\tilde{\sigma}^2 = \sigma^2$. The equality of the "implied" and "actual" parameters are testable restrictions.

For the no-arbitrage models the time series process of r_t itself has time varying parameters. But, assuming a constant price of risk, equation (4.11) implies that the deviation between the short rate and the lagged forward rate $r_t - f_{t-1,t}$ must be serially uncorrelated under the actual factor dynamics. The mean of $r_t - f_{t-1,t}$ under the actual probability measure is not determined by the model, and depends on the price of risk λ. As it is the innovation in the instantaneous spot rate, the variance must be equal to the conditional variance of r_t over a horizon of length h,

$$\mathrm{Var}(r_t - f_{t-1,t}) = \tilde{\sigma}^2 \left(\frac{1 - \exp(-2\tilde{\kappa}h)}{2\tilde{\kappa}} \right) \tag{4.15}$$

As in the equilibrium version of the model, consistency between the time series and cross sectional dimension implies $\tilde{\kappa} = \kappa$ and $\tilde{\sigma}^2 = \sigma^2$. However, since the variance of $r_t - f_{t-1,t}$ is the only moment that depends on these two parameters, they cannot be separately identified.

4.4 Why pooling?

Before developing a fully articulated econometric model, we briefly illustrate the problems in cross sectional estimation, and give the intuition why a panel data analysis will be helpful. For the example we consider the single factor Vasicek model.

Suppose the Vasicek model is true with parameters $\kappa = 0.04$, $\sigma = 2.5\%$ and $\theta = 0$. Figure 4.1 shows three yield curves starting with spot rates at 4%, 8% and 12% respectively. The yield curves are drawn for maturities between 0 and 10 years. For the selected

parameter configuration all three yield curves are upward sloping (only the curve starting at $r = 12\%$ has a slight hump) and almost parallel, although each of them must be downward sloping eventually, and converge to $Y_t(\infty) = \theta = 0$. Even though the Vasicek model implies that all yield curves converge to the same long term yield, this convergence is not visible over horizons up to 10 years when the mean reversion parameter κ is very small. A value of κ so close to the unit root is also plausible from a time series perspective (see for example Chan, Karolyi, Longstaff and Sanders (1992)).

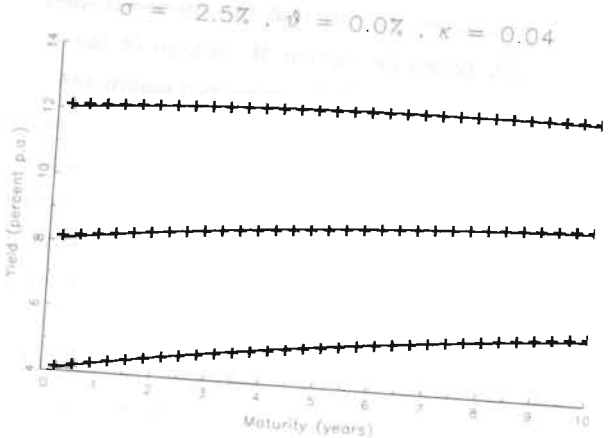

FIGURE 4.1: YIELD CURVE POOLING
The figure shows three yield curves generated by the Vasicek model with parameters $\theta = 0$, $\sigma = 2.5\%$ and $\kappa = 0.04$. The curves start at spot rates of $r = 4\%$, 8% and 12%, respectively.

In cross sectional estimation the parameters for the yield curve at time t are r_t, κ, σ, and θ. When estimated period by period, they can be different for each of the three curves. If the data are pooled, only r_t is different for the three curves, while κ, σ and θ are constant. In the example the yield curve data are without any measurement error, so that we can of course uniquely recover the true parameters of the data generating process. To show the problems with cross sectional estimators we simulate data by adding a small amount of measurement noise to the data, in the order of 8 basispoints at the short end to 3 basispoints at the long end. Errors with such a small standard deviation are not visible to the eye on the scale of figure 4.1.

In our Monte Carlo experiment parameters are estimated by nonlinear ordinary least squares. In effect, the Vasicek model is only nonlinear in κ. Conditional on κ, the estimator reduces to ordinary least squares with fixed regressors. The results in table 4.1 are based on

TABLE 4.1: MONTE CARLO RESULTS VASICEK YIELD CURVES

	RIMSE	r_1	r_2	r_3	κ	σ^2	θ
True values:							
	0.060	4	8	12	0.04	6.25	0
Fitted cross section models:							
Mean	0.001	4.00			0.066	4.33	0.92
Std.dev.	0.002	0.08			0.065	4.98	10.69
Minimum	0.000	3.68			-0.056	-42.14	-244.6
Maximum	0.020	4.28			0.336	7.97	7.96
Mean	0.001		8.00		0.100	1.80	1.35
Std.dev.	0.003		0.08		0.128	10.92	13.35
Minimum	0.000		7.77		-0.055	-149.6	-259.4
Maximum	0.033		8.27		0.682	10.44	10.55
Mean	0.004			12.00	0.201	-13.18	0.08
Std.dev.	0.008			0.08	0.351	45.41	36.42
Minimum	0.000			11.75	-0.360	-252.0	-1103.
Maximum	0.037			12.35	1.465	68.08	12.88
Pooled estimates:							
Mean	0.011	4.00	8.00	12.00	0.040	6.25	-0.29
Std.dev.	0.007	0.08	0.06	0.08	0.003	0.07	2.41
Minimum	0.000	3.67	7.82	11.73	0.028	6.02	-14.57
Maximum	0.038	4.25	8.20	12.36	0.054	6.56	6.17
True values:							
	0.200	4	8	12	0.04	6.25	0
Pooled estimates:							
Mean	0.053	4.00	8.00	12.01	0.040	6.19	-23.45
Std.dev.	0.039	0.40	0.31	0.39	0.016	0.62	100.11
Minimum	0.001	2.80	6.96	10.98	-0.013	-0.06	-1280.
Maximum	2.355	5.40	9.15	13.50	0.087	7.79	483.7

Notes: Yield curves are generated using the Vasicek model with the parameter values listed under "True values" including measurement noise:

$$\epsilon_t(\tau) = u_{1t} + \frac{\tau}{10}(u_{2t} - u_{1t})$$

where u_{1t} and u_{2t} are mutually uncorrelated normal random variables with zero mean. For the first part of the table the error standard deviations are 7 and 2 basispoints for u_{1t} and u_{2t} respectively. For the high noise case these standard deviations are 40 and 20 basispoints. All level parameters (r_{1t}, r_{2t}, r_{3t}, θ) are in units of percent per annum. The variance σ^2 is in percent per annum squared. The Root Integrated Mean Squared Error (RIMSE) is in basispoints. Parameters have been estimated by nonlinear least squares using 120 monthly spaced points on the yield curve. The longest maturity is ten years, the shortest one month.

small relative to both the cross sectional and the time series variation in yields. Otherwise, since the error deals with unmodelled phenomena, there are no apparent restrictions on its covariance structure.

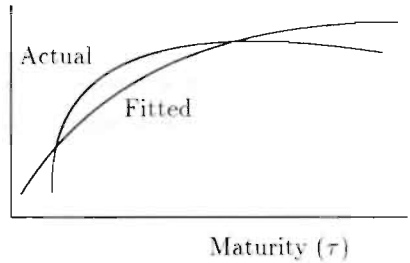

FIGURE 4.2: YIELD CURVE FITTING
The figure shows a hypothetical observed yield curve (Actual) and a yield curve implied by an equilibrium term structure model (Fitted). The observed yiled curve is drawn as a continuous line, because it is assumed that is constructed by interpolation methods like a spline function.

The term structure data are observed as a smooth function of maturity, to which we fit a curve that is also a smooth function of τ, as illustrated in figure 4.2. Therefore we assume that the correlation between error terms $\epsilon(\tau_i)$ and $\epsilon(\tau_j)$ depends on the distance between the terms to maturity:

$$\text{corr}(\epsilon_t(\tau_i), \epsilon_t(\tau_j)) = \phi^{|\tau_i - \tau_j|} \tag{4.17}$$

with $0 < \phi < 1$. Yields that are very close show high correlation, whereas yields that are far apart are less correlated. The specification resembles a cross sectional AR(1) error term.

Besides cross sectional correlation we also account for possible cross sectional heteroskedasticity of the error terms. The theoretical model can be written either in terms of log-prices or in yields. A transformation from yields to log prices $\ln P_t(\tau) = -\tau Y_t(\tau)$ induces maturity specific variances with a proportionality factor equal to τ^2. For that reason we specify the variance of the error term $\epsilon_t(\tau)$ as a function of τ:

$$\text{Var}(\epsilon_t(\tau)) = \omega^2 \tau^{-2d} \tag{4.18}$$

where ω^2 is a scale parameter, and d determines the sensitivity of the variance for the term to maturity. If $d = 0$ then the error terms are homoskedastic in a model for the yields. In

case $d = 1$, the model is homoskedastic in a regression model for (log-) bond prices. The heteroskedasticity implies a weighting scheme on the maturities: with $d = 0$ all yields have equal weights; with $d > 0$ more emphasis is put on long term yields. We estimate d along with the other parameters.

In matrix form the covariance structure for the cross sectional error terms is

$$\Sigma = \omega^2 S(\phi, d) \tag{4.19}$$

where S is a matrix of order $(N \times N)$ with typical element $s_{ij} = (\tau_i \tau_j)_*^{-d} \phi^{|\tau_i - \tau_j|}$.

As a further extension of the statistical model of the error term we allow for first order autocorrelation in the error terms of (4.16) like in Chen and Scott (1993),

$$e_t(\tau) = c e_{t-1}(\tau) + \epsilon_t(\tau) \tag{4.20}$$

where $\epsilon_t(\tau)$ is uncorrelated over time, and has the covariance structure specified in (4.19). We restrict the autocorrelation parameter c to be the same for each maturity τ in order to preserve the smoothness of the cross sectional error process.

The possible autocorrelation in the errors makes the equilibrium model flexible enough to approximate the extended model. If in the equilibrium model the autocorrelation parameter c approaches one, we obtain a model in first differences. In that case $B(\tau)$ is estimated from yield changes, while the level function $A(\tau)$ becomes unidentified. But empirically long maturity forward rates $F_{t-1,t}(\tau)$ are almost identical to the corresponding spot rates $Y_{t-1}(\tau)$. Yield changes $\Delta Y_t(\tau)$ will thus be very similar to the forward deviations $Y_t(\tau) - F_{t-1,t}(\tau)$ that are input to the extended models. Since $B(\tau)$ is identical in the equilibrium and extended models, an equilibrium model in first differences ($c = 1$) will be very similar to an extended model with $c = 0$, except at the short maturities.

Altogether the error specification consists of the parameters $\zeta = (\omega, \phi, d, c)$. The specification of the error term differs from other models in the literature in several respects. First, the error term has been added at the level of the discount price function, and not on the original traded individual (coupon) bond prices as in Brown and Dybvig (1986), Schotman (1996) and DeMunnik and Schotman (1994). Adding the error term at the level of discount bond prices has been proposed by Gouriéroux and Scaillet (1994), who motivate this choice from no-arbitrage conditions and invariance properties in modelling portfolios of bonds. However, Gouriéroux and Scaillet (1994) assume that the errors are uncorrelated across maturities, whereas our specification explicitly takes into account the smooth nature of observed yield curves.

Chen and Scott (1993) first eliminate the unobserved factors Z_t from (4.16), by using the exact relation between the factors and K different yields with maturities τ_{0j}, and then add

noise to the equations for all other maturities. This would imply that the particular yields with maturity τ_{0j} are always fitted exactly, *i.e.* $e_t(\tau_{0j}) = 0$ for $j = 1, \ldots, K$. Apart from this singularity, Chen and Scott (1993) allow for a fully unrestricted covariance matrix Σ. Bliss and Ritchken (1996) introduce the same singularity in their analysis of the extended Vasicek model, and also tightly parameterize the error covariance matrix. Frachot, Lesne and Renault (1995) and Bliss and Ritchken (1996) note that the choice of the pivotal maturities τ_{0j} is very influential for the empirical results. Therefore De Jong (1997) and Frachot, Lesne and Renault (1995) eliminate the unobserved state variables in a way that is invariant with respect to the maturity τ_{0j}. The number of unknown parameters in Σ in their specifications is of order N^2. Such a general unrestricted specification limits the number of maturities that can be included in the empirical analysis. In most of the empirical studies N is therefore only 4 or 5. Our tight parameterization of Σ allows N to be large, which helps to increase the power of the cross sectional tests. Eventually the strong positive correlation between neighboring maturities will put an upper bound on the information contents in the data.

Although tightly parameterized, our error covariance matrix is less restrictive than in other papers that use a broad spectrum of maturities. Gibbons and Ramaswamy (1993) assume absence of serial correlation in the measurement error. Many other studies (for example Geyer and Pichler (1997), Lund (1997)) assume absence of cross sectional correlation in the errors.

4.5.2 Estimation and Testing

The general model for both equilibrium and extended versions of the affine class is (4.16). For the estimation of the model standard panel data methods are applicable. The least restrictive assumption on the K factors Z_t is to treat them as a time series of unknown parameters, *i.e.* as fixed effects. The fixed effects approach is purely cross sectional; no time series information is used for the estimation of the structural parameters π of the model, and we therefore do not need assumptions on the dynamics of Z_t. Treating Z_t as parameters, the model in (4.16) can already be estimated on data for a single point in time t as in section 4.4. The panel data aspects arise from pooling the data for several cross sections.

We use Quasi Maximum Likelihood to estimate the structural parameters π, the error parameters ζ and the fixed effects $\{Z_t\}_{t=1}^{T}$. To handle the autocorrelation in the error

terms we adopt the following transformation for any time series variable X_t

$$\begin{aligned} X_t^* &= (1-c^2)^{\frac{1}{2}} X_1 && t=1 \\ &= X_t - cX_{t-1} && t=2,\ldots,T \end{aligned} \tag{4.21}$$

Using (4.21) on yields, factors, $A_t(\tau)$ and the errors we transform (4.20) to the model

$$Y_t^* = A_t^* + BZ_t^* + \epsilon_t, \tag{4.22}$$

for which we have the quasi-loglikelihood function

$$\begin{aligned} \ln L(\pi,\zeta,) &= -\tfrac{1}{2}NT\ln(\omega^2) - \tfrac{1}{2}T\ln|S| + \tfrac{1}{2}N\ln(1-c^2) + \\ &\quad -\frac{1}{2\omega^2}\sum_{t=1}^{T}(Y_t^* - A_t^* - BZ_t^*)'S^{-1}(Y_t^* - A_t^* - BZ_t^*), \end{aligned} \tag{4.23}$$

where $\mathbf{Z} = (Z_1,\ldots,Z_T)$. The likelihood can be concentrated analytically with respect to the fixed effects $\mathbf{Z}$ and the scale parameter ω^2. This is the crucial element of an affine model. For models not in the affine class the fixed effects cannot be concentrated out analytically, which would add T parameters to the non-linear numerical optimization of the likelihood function in (4.23), rendering these models intractable for analysis by conventional methods. For ω^2 the maximum likelihood estimator is given by

$$\hat{\omega}^2 = \frac{1}{NT}\sum_{t=1}^{T}(Y_t^* - A_t^*)'\mathbf{M}(Y_t^* - A_t^*)\,, \tag{4.24}$$

where $\mathbf{M} = S^{-1} - S^{-1}B(B'S^{-1}B)^{-1}B'S^{-1}$. Substitution of (4.24) in (4.23) gives the concentrated likelihood function

$$\ln L(\pi,c,d,\phi) = -\tfrac{1}{2}NT\ln\hat{\omega}^2 - \tfrac{1}{2}T\ln|S| + \tfrac{1}{2}N\ln(1-c^2) \tag{4.25}$$

which depends on only a small number of parameters, and can be easily maximized by numerical optimization routines.

The structural parameters only affect the first term in (4.25). It is instructive to consider this term in more detail. For the equilibrium models A_t does not depend on t. Let $\mathbf{V}$ be the sample covariance matrix of Y_t^* (which depends on c), and let $\bar{Y}$ be the sample mean of Y_t. Equation (4.24) can be rewritten as

$$\hat{\omega}^2 = \frac{1}{N}\left(\mathrm{tr}(\mathbf{MV}) + (\bar{Y} - A)'\mathbf{M}(\bar{Y} - A)\right)\,, \tag{4.26}$$

ignoring a term of order $(Y_T - Y_1)/T$. For the Gaussian model this decomposition implies that the function $A(\tau)$, and therefore the parameters θ and σ_j^2 are determined by the unconditional sample means $\bar{Y}$, *i.e.* the average shape of the yield curve. The function

$B(\tau)$ — and thus mean reversion parameters κ_j — are identified through all first and second moments of the transformed yields. QML defines a weighting scheme based on all possible $\frac{1}{2}N(N+3)$ moment conditions.

A similar decomposition for the extended Vasicek model in (4.11) shows that the volatility parameter is only determined through the convexity in $\bar{Y}-\bar{F}$ as a function of τ ($\bar{F}$ is the sample average of the N-vector with elements $F_t(\tau_i)$). Without using time series information all the information on σ^2 has to be extracted from the sample means of the deviation between the yield and the forward rate. Since this is not likely to be very informative data, estimates of σ^2 are presumably imprecise in the extended models.

The cross sectional specification of any affine term structure model is testable by comparing $A_t(\tau)$ and $B(\tau)$ with a less restrictive specification. To test for deviations from the theoretical model we augment the functional forms $A_t(\tau)$ and $B(\tau)$,

$$\begin{aligned} \tilde{A}_t(\tau) &= A_t(\tau)+\gamma' g(\tau) \\ \tilde{B}(\tau) &= B(\tau)+Dg(\tau)\,. \end{aligned} \tag{4.27}$$

where $g(\tau)$ is an L-vector of functions of τ, and γ and D are an L-vector and $(K\times L)$ matrix of parameters.

This tests considers the specification of the model keeping the number of factors constant. The number of factors itself can be examined by comparing with a model that has more factors. A formal test for the number of factors is not trivial due to the increasing number of incidental parameters in $\mathbf{Z}$. Every additional factor k introduces T new parameters in $\{Z_{kt}\}$, invalidating standard large T asymptotic inference. When the number of factors K equals the number of maturities N, the model will fit perfectly, and the likelihood function will go to infinity.

Unlike in standard factor models, or in principal components analysis, the "specific" risk terms $e_t(\tau)$ in (4.16) are not necessarily mutually uncorrelated. In fact, as discussed in section 4.2, they most likely are not, because of the smooth nature of the term structure data. Finding more than one factor in the covariance structure of interest rate data does not necessarily invalidate single factor term structure models. It is the shape of the factor loadings $B(\tau)$ that constitute the testable implications of the model.

4.6 Data

The data set consists of a panel of discount yields, which are constructed using the McCulloch (1975) spline procedure from US government bond data, available from the CRSP tapes. The data are monthly observations that span the period from January 1970 until

December 1994. At every time t the yield curve is represented by a cross section of $N = 16$ yields. The maturities are one through six months and one year until ten years.

Figure 4.3 shows the full data panel. Most striking is the predominance of parallel shifts of the yield curve. Apart from noisy behaviour at the short end, yield changes are almost horizontal. Yields with different maturities are therefore heavily correlated, both in levels and in deviation from lagged forward rates.

Summaries are presented in figure 4.4. The average yield curve in figure 4.4A is concave, as it should be in the Vasicek model in order to obtain a positive estimate of volatility. In deviation of the forward rate the concavity does not hold at the very short maturities of one- and two months in figure 4.4B.

Both for levels and changes the volatility in figures 4.4C and 4.4D decreases with maturity, as they should when interest rates are mean reverting. But the term structure of volatilities for the levels is flatter than for the changes. This is primarily due to the initial steep decrease in volatility for the very short term rates (maturities six months and less). Very short rates are different from the rest of the term structure. Volatility at longer maturities is still far from zero, indicating that the ten year interest rate is not a proxy for the infinite horizon yield that must be constant over time in the affine model. The observed yield curves have not converged to a single yield for long maturities.

The term structure of volatilities in figure 4.4D for first differences $Y_t(\tau) - Y_{t-1}(\tau)$ is almost identical to the volatility structure for the deviations between yields $Y_t(\tau)$ and lagged forward rates $F_{t-1,t}(\tau)$. This implies that the extended model is in effect a model of yield changes. This is brought out more clearly in the scatter diagram in figure 4.6.

Figure 4.5 shows yield curves ordered by the level of the ten year rate. For a one factor model, all curves with a common ten year rate should be identical apart from measurement error. Instead the curves spread out at the short maturities. One interpretation of these data is a second factor with strong mean reversion that only has an effect on short term yields. Another possibility is that measurement error for short maturities is relatively high.

4.7 Empirical results

In the empirical analysis we consider four different models: equilibrium and extended Gaussian models, with either one or two factors. Parameter estimates are reported in table 4.2. We first discuss the Vasicek single factor equilibrium model in detail, and then proceed with the other three models.

The parameter of interest in the Vasicek model is the implicit mean reversion κ, since it completely determines the volatility function $B(\tau)$. The panel model pools the yield

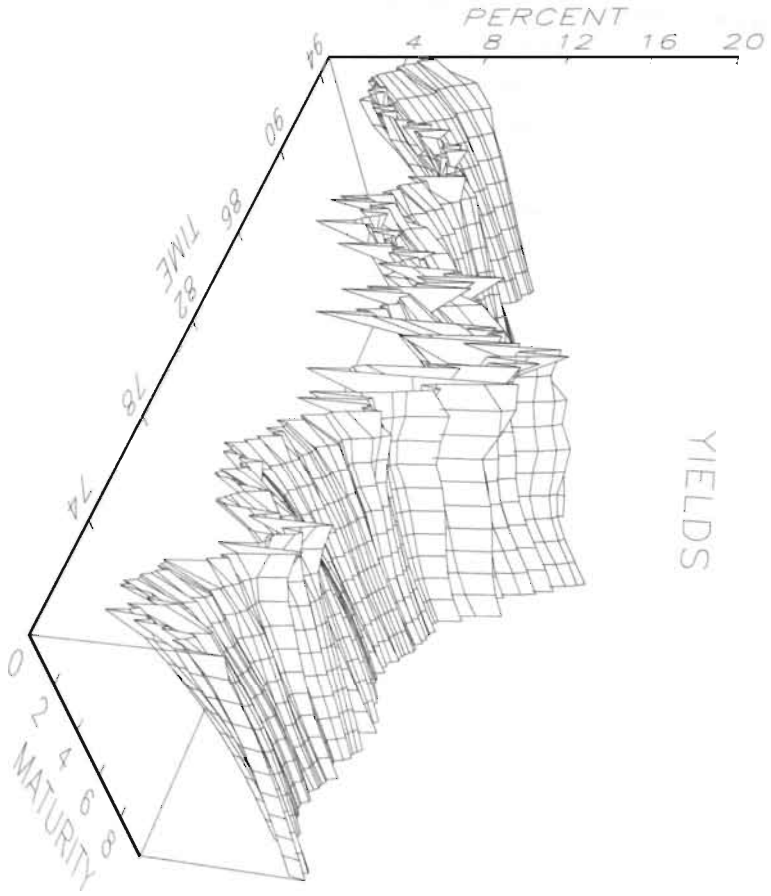

FIGURE 4.3: US INTEREST RATES

The figure shows monthly US discount yields for the period 1970 - 1994, and maturities between 0 and 10 years.

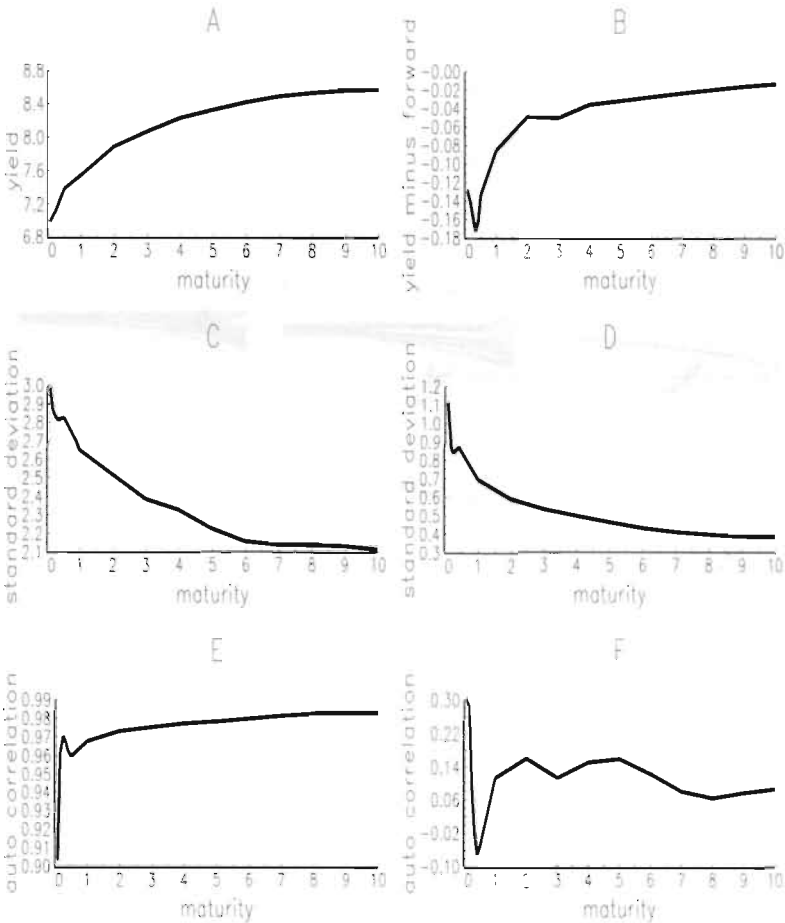

FIGURE 4.4: DATA SUMMARY

Panel A shows the sample mean of yields $Y_t(\tau)$ for different maturities, panel C the standard deviation, and panel E the first order autocorrelation. Yields are measured in percent per annum. Panels B, D en F provide the same summary for yields in deviation of the lagged forward rate $F_{t-1,t}(\tau)$ (solid line) and first differences of yields (dashed line). All data are monthly observations for the period 1970-1994.

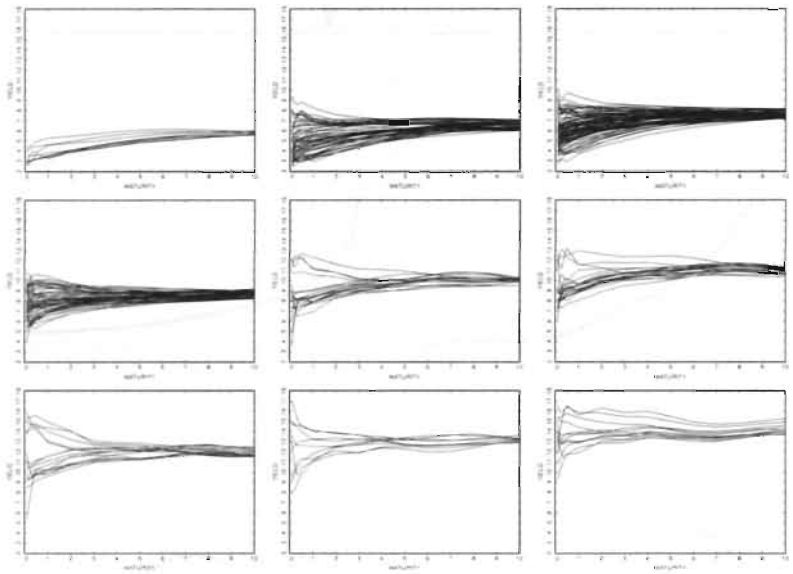

FIGURE 4.5: A SECOND FACTOR

The figure shows observed yield curves ordered with respect to the level of the ten year discount rate.

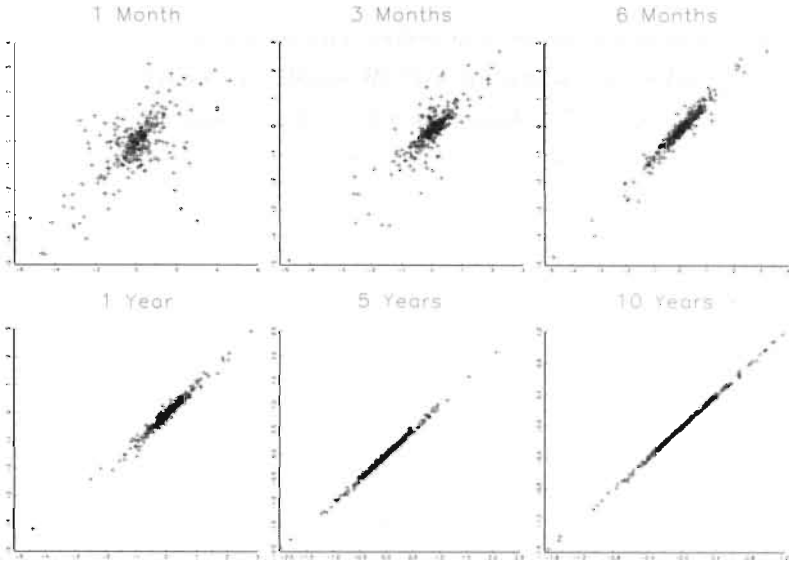

FIGURE 4.6: DISCOUNT YIELDS AND FORWARD RATES

The figure shows scatter plots with $Y_t(\tau) - F_{t,t-1}(\tau)$ on the vertical axis and $Y_t(\tau) - Y_{t-1}(\tau)$ on the horizontal axis for selected values of τ.

curves for many different months. During the sample period short term interest rates have fluctuated between 5 and 15 percent, *i.e.*, the yield curves start from very different levels of the instantaneous spot rate. When all these yield curves are forced to converge to a single infinite yield, the only way that any single factor model can achieve this, is by setting the mean reversion parameter at a very small value. The point estimate implies a monthly autoregressive coefficient of 0.997, or equivalently a half-life of shocks equal to $\ln 2/\kappa = 17$ years. From the cross sectional perspective interest rates have to be near integrated series. Although the mean reversion is low, it is estimated very precisely and is significantly different from zero, with a t-statistic that is larger than 8.

At the very long end of the yield curve, the estimate of the infinite yield θ is negative and imprecise. It appears that this parameter is not identified in the data. The poor estimates are related to the low mean reversion, since the constant infinite maturity yield becomes unidentified when κ is close to zero. We simply have no reliable data about very long term interest rates.[10,11,12] Although the infinite yield is negative, the model generates upward sloping, almost parallel, yield curves for maturities up to 10 years.

Due to the low mean reversion the term structures are far from convergence to their common infinite horizon yield at a maturity of ten years. It might be considered a serious drawback of the Vasicek model, like all other one factor models, that it implies one constant infinite yield. But even in a multifactor world, it is not trivial to construct an equilibrium model without this feature. Variability of long maturity interest rates will always imply low mean reversion.[13]

Other panel data studies, referred to in the introduction, also incorporate time series information. Instead of estimating the factor as an unknown parameter, they treat it as a latent variable that follows the Gaussian process (4.14). The parameters κ and σ appear in both the cross sectional as well as the time series dimension of the model, in other words in both the risk neutral and actual dynamics. That way these studies use all possible information in the data at the cost of making the assumption of a constant price of risk. But the additional information in the time series seems weak. When we use the Kalman filter to estimate the model by maximum likelihood with r_t as a latent

[10] The CRSP data set does contain bonds with a maturity of 30 years. We included these bonds, when we estimated spline functions to create discount yields. However, the spline function shows very large standard errors at maturities beyond ten years, indicating that the yield curve data are measured with much error for these long maturities. For that reason we did not extend our discount yield data beyond the ten year maturity.

[11] The CIR model yields similar results. Although mean reversion is even lower at 1.25e-6 and the infinite yield is positive at 6.35 by construction, the fit of the term structure is almost identical.

[12] Setting $\theta = 0$ instead of the ridiculously large negative value makes absolutely no difference for the estimate of κ. It only has some effects on σ.

[13] See for example the theoretical discussion in El Karoui, Frachot and Geman (1998).

TABLE 4.2: PARAMETER ESTIMATES

	Equilibrium			Extended		
	One Factor	Two factors		One Factor	Two factors	
κ	0.042	0.003	0.311	0.095	0.012	0.081
	(0.005)	(0.001)	(0.010)	(0.002)	(0.004)	(0.013)
σ	0.321	0.290	0.797	0.507	0.210	0.189
	(0.023)	(1.122)	(0.248)	(0.026)	(0.014)	(0.015)
θ	-8.74	-4539				
	(5.18)	(1003)				
ϕ	0.725	0.031		0.720	0.503	
	(0.026)	(0.045)		(0.001)	(0.005)	
d	0.763	0.746		0.748	0.794	
	(0.008)	(0.046)		(0.003)	(0.008)	
c	0.785	0.693		-0.314	-0.356	
	(0.024)	(0.004)		(0.005)	(0.027)	
ω	0.408	0.258		0.398	0.265	
$\ln L$	31364	32214		31305	32143	

Notes: The table reports the estimation results for four versions of the (extended) Vasicek model in a panel with fixed effects. The first half of the columns refer to the equilibrium model, the last half of the columns to the extended Vasicek model. The parameters κ, σ and θ denote the mean reversion, volatility and infinite yield. The parameters ω, ϕ and d define the covariance matrix of the cross sectional error term; c is the autocorrelation parameter in the error term. Time in measured in years, so that for example $\ln 2/\kappa$ measures the halflife in years. Other parameters, like θ and σ are converted to percent per year. Standard errors are in parentheses. $\ln L$ denotes the log likelihood.

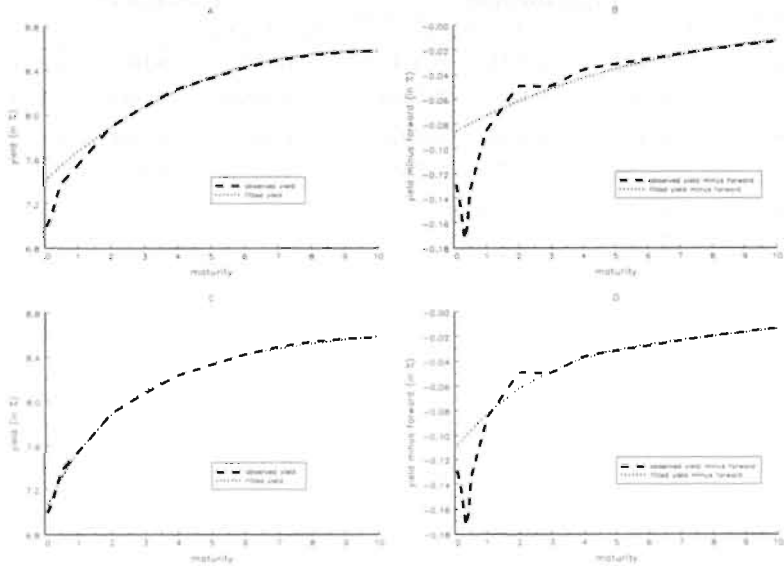

FIGURE 4.7: AVERAGE YIELD CURVE

This figure shows the average yield curve along with the fitted average yield curve for four different models: One factor equilibrium Vasicek (A), One factor extended Vasicek (B), Two factor equilibrium Vasicek (C), and Two factor extended Vasicek (D).

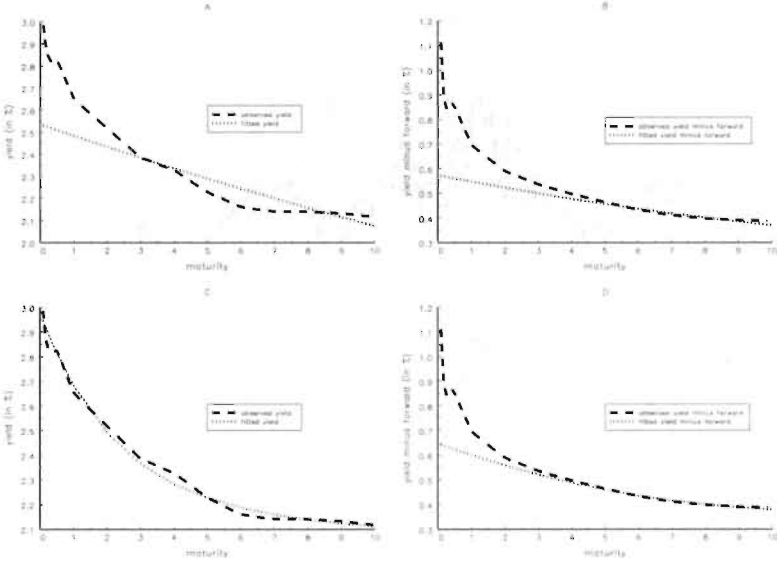

FIGURE 4.8: TERM STRUCTURE OF VOLATILITIES

This figure shows the standard deviations of yields with different maturities along with the fitted standard deviations for four different models: One factor equilibrium Vasicek (A), One factor extended Vasicek (B), Two factor equilibrium Vasicek (C), and Two factor extended Vasicek (D). The fitted strandard deviation refer to the volatility of $B(\tau)z_t$. Units are percent per annum.

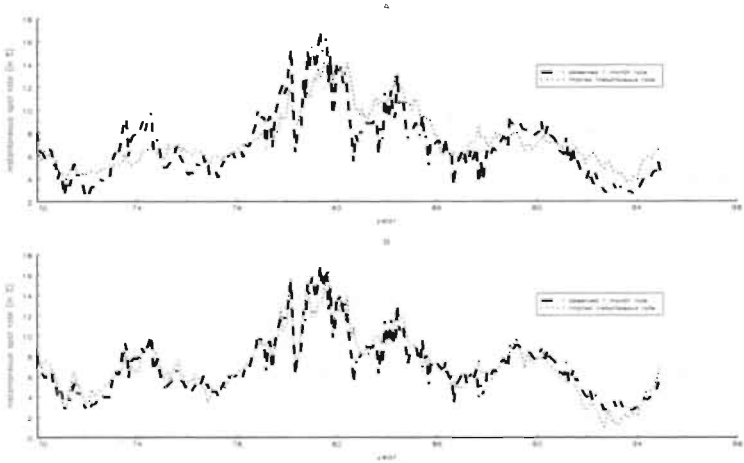

FIGURE 4.9: IMPLIED SPOT RATES

The figure shows the observed one month discount yield and the estimated spot rates r_t for the equilinbrium Vasicek model. Panel A refers to a one factor model, and panel B to the two factor model, where R_t is the sum of the two factors.

variable, the parameter estimates remain unchanged. Our point estimates are $\hat{\kappa} = 0.039$ and $\hat{\sigma} = 0.337$ with standard errors that are only slightly less than those in table 4.2. The cross sectional information dominates the time series information in the data. This should not come as a surprise, since time series data are known not to be informative on mean reversion. That was one of the conclusions of Chan, Karolyi, Longstaff and Sanders (1992). When time series and cross sectional parameters are allowed to differ in the way discussed below (4.14), we indeed find that the time series mean reversion $\hat{\kappa} = 0.21$ gets a standard error of 0.12, and is not significantly different from zero.

The diagnostic tests give many indications that the one factor model is misspecified. First, the likelihood ratio test rejects the structure of $A(\tau)$ and $B(\tau)$ imposed by the Vasicek model. The deviations are illustrated in figures 4.7A and 4.8A. Figure 4.7 shows the average observed term structure and the average of the estimated term structure, given by $A(\tau) + B(\tau)\bar{z}$, where $\bar{z}$ is the sample mean of the estimated factor. On average the actual term structure is steeper than what would be consistent with the one factor Vasicek model. Figure 4.8 shows the volatility structure. The estimated volatility is computed as the standard deviation of $B(\tau)z_t$. Short term rates are much more volatile than implied by the model. Another way to interpret figure 4.8 is as a measure of fit. The ratio of fitted to actual variance, $\mathrm{Var}(B(\tau)z_t)/\mathrm{Var}(Y_t(\tau))$ for different maturities can be used as an R^2 measure of goodness of fit. The explained variance quickly rises from 31% for the one month interest rate to over 99% for yields with maturities of 3 years and longer.

Second, the estimate of the cross sectional heteroskedasticity parameter d implies that parameter estimates are mostly determined by data from the long end of the maturity spectrum, and that the model does not fit the short end at all. With $\hat{d} = 0.75$, the relative weight of the ten year yield compared to the one month rate is $(\frac{10}{1/12})^{\hat{d}} \approx 36$, against equal weighting when $d = 0$.

Third, the estimate of ϕ states that cross sectional dependence of the errors is strong. For yields that differ only one month in time to maturity, the correlation coefficients are about $\phi^{1/12} = 0.97$. There is a strong common component left in the errors. This common component also has strong autocorrelation, but is far away from the unit root.

Fourth, figure 4.9 depicts the implied instantaneous spot rate, *i.e.*, the estimated fixed effect $\hat{r}_t$, together with the observed one month rate (as a proxy for the short rate). The implied rate is less volatile than the observed rate. It looks as if the actual one month interest rate contains a lot of transitory noise. The same transitory noise is also visible in the yield curve changes in figure 4.3.

All these four empirical diagnostics point at a second factor, which has stronger mean reversion than the dominant first factor, and which therefore will mostly affect the shorter

maturities. The second column in table 4.2 show the estimation results. As anticipated we find one factor with very low mean reversion, like in the one factor case. The second factor is much more volatile but strongly mean reverting. This model fits the data much better. The restrictions on $A(\tau)$ and $B(\tau)$ can not be rejected, and therefore the actual and fitted first and second moments match almost perfectly in figures 4.7 and 4.8. The two factors account for more than 95% of all variation in interest rates. The remaining error terms are small. Moreover, most of the cross sectional correlation in the errors has disappeared.

The high mean reversion of the second factor is entirely consistent with the time series behavior of short term interest rates. Both Andersen and Lund (1996) and Balduzzi, Das and Foresi (1998) find that the short term interest rate can be described as a time series that quickly reverts towards a slowly changing mean. Dai and Singleton (1997) show that such a model, which they call a "cascade" model, is just an alternative parameterization of the two-factor model.

A disturbing diagnostic is that the residual autocorrelation remains. The second factor only fixes the misspecification at the short end, but does not affect the longer term yields. Since most of the autocorrelation comes from the longer term yields, the second factor can not take that away.[14] When the theoretical yield curve does not match the observed yield curve, the error persists for some time. That is exactly the type of behavior that can be modelled by an extended model. An extended model starts from an exogenously given yield curve, and assumes that movements away from the initial conditions satisfy a no-arbitrage constraint. If the current yield curve is not consistent with an equilibrium model, then the error will only partly be corrected in next period's yield curve. When an equilibrium is fitted to data from an extended model with an arbitrary initial condition, the error terms will exhibit strong positive autocorrelation. Estimation of an extended model is therefore directly motivated by the results from the equilibrium model.

The estimate of the mean reversion in the extended Vasicek model in table 4.2 is close to the estimate for the equilibrium version of the model. Like the equilibrium model, the one factor extended Vasicek model has problems fitting the short term interest rate movements. The observed volatility is much higher than the volatility $\mathrm{Var}(B(\tau)(r_t - f_{t-1,t}))^{\frac{1}{2}}$ implied by the estimated factor.

In the two factor extended model, the first factor is almost a random walk. Since the equilibrium and extended Vasicek models are not nested, they cannot be formally tested against each other. The two models aim at fitting different sets of moments in the data. Still the likelihood value of the extended two-factor model falls short of the likelihood value

[14] This of course means that our parameterisation of the measurement error structure is too restrictive, since it imposes the same autocorrelation for all maturities. But this does not affect the consistency of the parameter estimator, and only means that a more efficient estimator exists.

of the equilibrium model. Other diagnostics also suggest that the equilibrium version of the model fits the data better.

A major difference between the extended and equilibrium models is the autocorrelation in the residuals. First, for the equilibrium model, specified in levels, the estimated autocorrelation is about $c \approx 0.75$. For the extended Vasicek model, which is practically equivalent to taking first differences of the yields, the autocorrelation coefficient becomes negative at about $c \approx -0.30$. Negative autocorrelation is an indication that working with monthly changes in yields leads to 'overdifferencing'. The extended Vasicek model calibrates too often! What is missing in the extended Vasicek model is an "error correction" factor that measures the deviation of the current term structure from the expected yield curve based on an equilibrium model. The levels of interest rates contain useful information about the parameters that is ignored in the extended model. This could explain why parameter estimates are more precise for the equilibrium version of the model.

Second, the estimated volatility structure in figure 4.8 indicates that the two-factor model still does not account for the relatively high volatility at the short end of the term structure. This is directly related to the low mean reversion in the second factor. The low value for $\hat{\kappa}_2$ arises from the covariance structure of $Y_t(\tau) - F_{t-1,t}(\tau)$. The correlations between $Y_t(\tau) - F_{t-1,t}(\tau)$ are generally lower than between the yield levels $Y_t(\tau)$, even after correcting for autocorrelation as in $Y_t^*(\tau)$. The data need the second factor to improve the fit of movements in long term rates.

4.8 Conclusions

We have proposed a fixed effects panel data model for the term structure of interest rates. This framework allows us to estimate the risk neutral process of the factors without making additional assumptions on the actual time series behavior of the factors. In this framework we treat single and multifactor models as well as equilibrium and extended term structure models. Two issues are important in the econometric model. First, we consider the affine class of term structure models, since this class allows the use of linear panel data methods. Second, the model takes into account the natural ordering of interest rates in the maturity dimension to parameterize the covariance matrix of the error terms.

Pooling results in very sharp point estimates of the structural parameters. Most salient is the implied mean reversion in a one factor Gaussian model, which is very low but statistically different from zero. Such precise estimates are not available from time series analysis of 25 years of monthly data. Moreover the equilibrium version model fits very tight for maturities larger than one year. At the very short side of the yield curve, we find a lot

of transitory noise, which requires a second factor. For the extended Gaussian models we find strong negative autocorrelation as evidence of overdifferencing. The extended models ignore useful information in the levels of the yield curve.

Chapter 5

An Empirical Investigation to the Longstaff-Schwartz Term Structure Model

5.1 Introduction

In the previous chapter an econometric framework has been set up to apply panel data techniques for affine term structure models. In that particular panel data model only cross sectional information was incorporated and no reference was made to the time series dynamics of the underlying factors. From the time series literature it is well-known that interest rate volatility is time varying, which suggests to include it as a second factor in the term structure model. In this chapter we propose to use a panel data approach to test the Longstaff-Schwartz (1992) term structure model, since it is a tractable model. Two issues are important. First, in the LS model volatility of the spot rate is included as a second factor next to the spot rate itself, and hence we compare it with the nested CIR 1 factor model to find out whether volatility is an important second factor. Second, we compare the term structure implications of the panel data approach with the implications of a time series approach on the underlying factors. Because these estimation approaches focus on different moments in the data their implications for the term structure or for other derivative securities need not be the same.

A vast part of the literature deals with specifications of single or multifactor models, where the aim is to find a model that best describes the time series dynamics of the underlying factors. In the so-called single factor models, the instantaneous spot rate is adopted as the factor that drives the yield curve. Examples are the Vasicek (1977) and the Cox, Ingersoll and Ross (1985b, CIR) model. Chan, Karolyi, Longstaff and Sanders

(1992, CKLS) provide an overview of different specifications for the spot rate, in which they include various parameter values to take account of the level effect between spot rate and volatility.

Single factor models are misspecified for two reasons. They do not fit the dynamics of the underlying factor, usually the short term interest rate, and they do not match the shape of the yield curve. Andersen and Lund (1997) extend the specification of the spot rate process by including a stochastic volatility process and level effect to better represent the dynamics of the spot rate. Dybvig (1989) suggests that volatility is an important second factor for modeling both the yield curve and options on bonds. Longstaff and Schwartz (1992a) adopt square root processes for both the instantaneous spot rate and the associated volatility.[1]

The inclusion of volatility as a second factor as proposed by Andersen and Lund (1997) provides an adequate description of the first and second moments of the instantaneous spot rate. Hence the underlying factor dynamics are well described. This does however not necessarily imply that the associated yield curve is adequately described as well. First, in the time series process for the instantaneous spot rate the parameters are expressed in terms of the actual probability measure, whereas the yield curve depends on parameters under the risk adjusted probability measure. To obtain the parameters under the risk adjusted probability measure we require the market price of risk, which is not available from time series data only. So, for purposes of yield curve modeling a time series estimation approach of the spot rate is not sufficient and adoption of cross sectional information is required. Second, a good fit of the yield curve is not guaranteed because in time series approaches no restrictions have been imposed to fit the yield curve.

The apparent success of stochastic volatility as a second factor in time series approaches, hence, motivates an investigation towards the cross sectional properties of this type of two factor model. Apart from the fact that panel data estimation deals with the parameters under the risk adjusted probability measure directly, there exist other good reasons to consider a panel data approach. First, since we are interested in the implications for the yield curve, it makes intuitively sense to match the moments of yield curve data. Second, using the entire set of yield curves simultaneously in a panel data model, will lead to more efficient estimates in comparison with the use of time series data for the underlying factors only. Third, in a time series approach the underlying factors are often not directly

[1]Other two factor models have been suggested by Brennan and Schwartz (1979) who use the long-term interest rate as the second factor. In Schaefer and Schwartz (1984) the spread between the long-term and short-term interest rate is adopted as a second factor. CIR (1985b) use an exogenously specified process for uncertain inflation in addition to the short term interest rate. Balduzzi, Das, Foresi and Sundaresam (1996) introduce the concept of a time varying mean which they call the central tendency.

observable and have to be approximated, which may lead to biases. In the panel data framework, only yield curve data are required, which are observable and the unobserved factors are treated as incidental parameters, as proposed by Brown and Dybvig (1986), DeMunnik and Schotman (1994) and Bams and Schotman (1997). Treating the time series of factors as unknown parameters leads to a panel data model with fixed effects.

Here, we test the cross sectional properties of the Longstaff and Schwartz (1992a, LS) model. The LS model is attractive since it adopts the same factors, spot rate and spot rate volatility, as Andersen and Lund (1997) adopt in their time series approach. Moreover, the model belongs to the affine class of term structure models which provides a closed form solution for the yield curve.[2] This enables the use of standard panel data techniques, which makes the model attractive for a cross sectional analysis, as in chapter 4.

In LS a time series approach is pursued to estimate the LS model and cross sectional restrictions have been tested by use of the GMM approach, applied on changes of yields. They cannot reject the overidentifying cross sectional restrictions. However, since the model is estimated and tested on changes in yields not all parameters are estimated and implications for the level of yields cannot be directly inferred. In the panel data approach we focus on the first and second unconditional moments of the yields as well as on the covariance matrix of the yields to estimate the LS model.

In the next section we describe the Longstaff-Schwartz model, in section 3 we discuss the econometric specification and the estimators. Section 4 presents summary statistics of the data. In section 5 we presents results, where we compare the term structure implications for the CIR 1 factor model, for the LS model both estimated from a time series perspective as proposed by LS, and from a cross sectional perspective as proposed in chapter 4. To investigate whether the alternative approaches are really different, we also compare the implications for cap pricing for all approaches. In section 6 we finish the chapter with a number of concluding remarks.

5.2 Model specification

LS derive a term structure model in the equilibrium approach pioneered by CIR (1985b). In the equilibrium approach an economic framework is set up that leads to the fundamental partial differential equation.[3] This PDE holds for all securities that are derivatives of the underlying factors. The term structure of interest rates is a derivative of the underlying factors, and so are all other interest rate derivative securities. The PDE includes the

[2]The affine class of term structure models has been studied in detail in Duffie and Kan (1997).

[3]See Duffie (1996) for a general treatment of term structure theory

parameters of the underlying factor process under the risk neutral probability measure. Hence for term structure modeling we are only concerned with parameters under this measure.[4]

We consider the two factor LS model, since this enables formal interpretation of each factor in terms of economic variables, spot rate and volatility of the spot rate. The LS model is closely related to the two factor CIR model, which belongs to the affine class of term structure models. For affine term structure models closed form solutions for the yield to maturity are available, which makes the LS model analytically tractable.

LS start with the introduction of two unspecified positive state variables that follow uncorrelated mean-reverting square root processes of the CIR type

$$dX_1 = (C_1X_1 + d_1)dt + \sqrt{\beta_{11}X_1}dW_1 \tag{5.1}$$

$$dX_2 = (C_2X_2 + d_2)dt + \sqrt{\beta_{22}X_2}dW_2 \tag{5.2}$$

where X_1 and X_2 are the state variables, W_1 and W_2 are uncorrelated Brownian motions, C_1 and C_2 denote the mean-reversion parameters, d_1 and d_2 are scalars and β_{11} and β_{22} determine the scale of the variances. These processes are both under the actual probability measure. For interpretation purposes, LS transform these unspecified state variables to two factors which are of economic interest, the spot rate and the volatility of the spot rate. From Duffie and Kan (1996) a well-known property of affine term structure models, with uncorrelated state variables, is that the sum of the state variables can be normalized to the instantaneous spot rate, hence

$$r = X_1 + X_2 \tag{5.3}$$

where r denotes the instantaneous spot rate. Because the state variables are uncorrelated, the variance of the spot rate follows directly as

$$V = \beta_{11}X_1 + \beta_{22}X_2 \tag{5.4}$$

where V denotes the variance of changes in the instantaneous spot rate. The linear relationships between the unknown state variables and the factors in equations (5.3) and (5.4) imply that the LS term structure model can also be expressed in terms of two other factors, spot rate and volatility of the spot rate, instead of the unknown state variables. The dynamics of r and V follow directly by solving (5.3) and (5.4). Both r and V follow square-root processes

$$dr = \left[(d_1 + d_2) + \left(\frac{C_1\beta_{22} - C_2\beta_{11}}{\beta_{22} - \beta_{11}}\right)r + \left(\frac{C_2 - C_1}{\beta_{22} - \beta_{11}}\right)V\right]dt \tag{5.5}$$

[4]See chapter one for an exposition of the factor specification under the actual probability measure, the risk neutral measure and the relation between these two.

$$+\beta_{11}\sqrt{\frac{\beta_{22}r - V}{\beta_{11}(\beta_{22} - \beta_{11})}}dW_1 + \beta_{22}\sqrt{\frac{V - \beta_{11}r}{\beta_{22}(\beta_{22} - \beta_{11})}}dW_2$$

and

$$dV = \left[(\beta_{11}d_1 + \beta_{22}d_2) - \left(\frac{\beta_{11}\beta_{22}(C_2 - C_1)}{(\beta_{22} - \beta_{11})}\right)r - \left(\frac{\beta_{22}C_2 - \beta_{11}C_1}{(\beta_{22} - \beta_{11})}\right)V\right]dt \quad (5.6)$$
$$+\beta_{11}^2\sqrt{\frac{\beta_{22}r - V}{\beta_{11}(\beta_{22} - \beta_{11})}}dW_1 + \beta_{22}^2\sqrt{\frac{V - \beta_{11}r}{\beta_{22}(\beta_{22} - \beta_{11})}}dW_2$$

Berardi and Esposito (1997) introduce the term 'shape-based' inversion of affine term structure models, in case the volatility of the spot rate is included in a transformation of the basic affine term structure model, given in equations (5.1) and (5.2). They show that the volatility of the short rate measures the convexity of the forward rate curve, which explains the term.

As opposed to the processes for the original state variables, the processes for r and V will be correlated, and it can be shown that this correlation is positive. In the literature the positive relation between interest rate levels and the underlying interest rate process is known as the level effect and reported in, for example, CKLS, Koedijk, Nissen, Schotman and Wolff (1997, KNSW), Andersen and Lund (1997) and others. From equations (5.3) and (5.4) and the fact that both state variables, X_1 and X_2, and the transformed factors, r and V, are positive, the following relation between r and V follows

$$\begin{cases} \beta_{11}r < V < \beta_{22}r & \beta_{22} > \beta_{11} \\ \beta_{22}r < V < \beta_{11}r & \beta_{22} < \beta_{11} \end{cases} \quad (5.7)$$

This relation takes account of the level effect between spot rate and volatility, but it is flexible in a sense that it does not assume an exact one-to-one relation between volatility and spot rate. This is different from the specifications in CKLS, where the variance is proportional to the level of the spot rate and therefore completely determined by the level of the spot rate.

Because the underlying state variables are affine, the specification for the yield curve is also affine. For the affine class of term structure models with independent factors, an analytical expression for the yield curve is available. See, for example, Duffie and Kan (1996), Berardi and Esposito (1997) or Frachot and Lesne (1993) for a derivation. LS (1992a) express the yield curve in terms of the spot rate and the volatility of the spot rate, instead of the independent factors. The model is still affine in the factors and is of the form

$$Y_t(\tau) = A(\tau) + B(\tau)f_t \quad (5.8)$$

where $Y_t(\tau)$ represents the yield at time t with time-to-maturity τ, $A(\tau)$ is a maturity dependent scalar, $B(\tau)$ is a (1×2) vector of factor loadings, f_t is the (2×1) vector of

factors, including spot rate and spot rate volatility at time t

$$f_t = \begin{pmatrix} r_t \\ V_t \end{pmatrix}$$

The model in equation (5.8) relates the yield with maturity τ to a constant and two factors. The relationship between the yield and the factors is linear which makes the factor loadings readily interpretable as the effect of a change in the factor on the yield curve. The linearity also makes the model attractive from an econometric point of view as we will elaborate on in the next section.

The functional form of $A(\tau)$ and $B(\tau)$ depends upon the parameters of the factor processes, possibly corrected with the market price of risk. In the LS model the market price of risk is proportional to factor X_1. This results in a market price of risk of the form λX_1, where λ is a proportionality constant. The functional specification of the scalar for maturity τ, is given by

$$A(\tau) = -\frac{d_1}{\beta_{11}}\left[\xi - C_1 + \frac{2}{\tau}\ln D_1(\tau)\right] - \frac{d_2}{\beta_{22}}\left[\psi - C_2 + \lambda + \frac{2}{\tau}\ln D_2(\tau)\right] \tag{5.9}$$

Let $B(\tau) = [B_1(\tau)\ \ B_2(\tau)]$, then the factor loading associated with the instantaneous spot rate is of the form

$$B_1(\tau) = -\frac{\beta_{11}\xi\left[\exp(\psi\tau) - 1\right]D_2(\tau) - \beta_{22}\psi\left[\exp(\xi\tau) - 1\right]D_1(\tau)}{\tau\xi\psi(\beta_{22} - \beta_{11})} \tag{5.10}$$

and the factor loading for the variance of the instantaneous spot rate is given by

$$B_2(\tau) = -\frac{\psi\left[\exp(\xi\tau) - 1\right]D_1(\tau) - \xi\left[\exp(\psi\tau) - 1\right]D_2(\tau)}{\tau\xi\psi(\beta_{22} - \beta_{11})} \tag{5.11}$$

where

$$\xi = \sqrt{2\beta_{11} + (\lambda - C_1)^2} \tag{5.12}$$

$$\psi = \sqrt{2\beta_{22} + C_2^2} \tag{5.13}$$

$$D_1(\tau) = \frac{2\xi}{(\xi - C_1 + \lambda)\left[\exp(\xi\tau) - 1\right] + 2\xi} \tag{5.14}$$

$$D_2(\tau) = \frac{2\psi}{(\psi - C_2)\left[\exp(\psi\tau) - 1\right] + 2\psi} \tag{5.15}$$

The LS model is a transformed version of the CIR two-factor model. A special case of the LS model is the CIR one-factor model, where the instantaneous spot rate, r, is the only factor driving the yield curve with the following dynamics

$$dr = (C_1 r + d_1)\,dt + \sqrt{\beta_{11} r}\,dW \tag{5.16}$$

The same diffusion process holds for the second state variable in the LS model. Hence, the CIR model is nested in the LS model, by imposing the restrictions $d_2 = C_2 = \beta_{22} = 0$ and rederiving the term structure of interest rates. The yield curve representation for the CIR model is given by

$$Y_t(\tau) = A^*(\tau) + B^*(\tau) r_t \tag{5.17}$$

where $A^*(\tau)$ is a restricted version of $A(\tau)$ in the LS model

$$A^*(\tau) = -\frac{d_1}{\beta_{11}} \left[\psi - C_1 + \lambda + \frac{2}{\tau} \ln D_1(\tau) \right] \tag{5.18}$$

and $B^*(\tau)$ is a restricted form of $B_1(\tau)$ in the LS model

$$B^*(\tau) = \frac{2\left(e^{\xi\tau} - 1\right)}{\left[(\xi - C_1 + \lambda)\left(e^{\xi\tau} - 1\right) + 2\xi\right]\tau} \tag{5.19}$$

Testing the LS model against the CIR model, hence, boils down to testing whether d_2, C_2, and β_{22} are actually different from zero in the LS model. In that case, the LS model is an improvement over the CIR one-factor model, which means that the volatility of the spot rate is a factor that does contribute to the behaviour of the yield curve.

Because the yield curve depends upon both the short-term rate of interest and the associated volatility process, the model is able to represent a greater variety of shapes like humped yield curves. Since the volatility factor can be changed while the spot rate is fixed, the LS model is also able to model term structures that are less flat and that capture the spread between longer term yields and the short rate. This is a point where single factor models tend to fail. In a single factor model the parallel shifts at the long end of the yield curve force the spot rate to exhibit extreme low mean reversion, which is inconsistent with the higher mean reversion we observe in the time series dynamics of spot rates.

In the empirical application we will consider the in-sample performance of the term structure models, i.e. how do term structure models fit the yield curve. We are also interested in their out-of-sample behaviour and that is why we investigate the implications of interest rate cap pricing for the alternative models. A simple interest rate cap, is a derivative security that pays the difference between the short rate and the cap rate, c, at maturity. Let $\mathrm{CAP}_t(X_1, X_2, c, \tau)$ denote the value of an interest rate cap at time t, with cap rate c, and time-to-maturity τ. A cap depends upon the underlying state variables X_1 and X_2 and its current value is found by applying the risk neutral valuation principle

$$\mathrm{CAP}_t(X_1, X_2, c, \tau) = E^Q\left[\max(X_1 + X_2 - c, 0)\right] \exp\left[-\tau Y_t(\tau)\right] \tag{5.20}$$

where $E^Q[\cdot]$ denotes the expectation operator under the risk neutral probability measure. Under this measure the diffusion processes for X_1 and X_2 are given by the dynamics

$$dX_1 = \left(\left[C_1 - (\xi + C_1)(1 - D_1(\tau))\right] X_1 + d_1\right) dt + \sqrt{\beta_{11} X_1} dZ_1 \tag{5.21}$$

and

$$dX_2 = \left(\left[C_2 - (\psi + C_2 - \lambda)(1 - D_2(\tau)) \right] X_2 + d_2 \right) dt + \sqrt{\beta_{22} X_2} dZ_2 \qquad (5.22)$$

where Z_1 and Z_2 denote Brownian motions and $D_1(\tau)$ and $D_2(\tau)$ are given in equations (5.14) and (5.15), respectively. For the simpler CIR one-factor model a similar approach is applicable.

5.3 Econometric specification

Different estimation methods exist to estimate the LS term structure model. These methods differ with respect to the type of data they require and with respect to the moments of the data the model is fitted on. Generally speaking, in term structure models the focus can be on the underlying factors, on the yield curve, or on a combination of these two. Time series estimation approaches of yield curve models focus on the dynamics of the underlying factors as in equations (5.5) and (5.6), whereas in cross sectional approaches moments in the yield curve determine the estimation results, as in equation (5.8). In a panel data approach the model incorporates all cross sections simultaneously.

The underlying economic framework would imply that both approaches are consistent with regard to the implications for the term structure. However, since both approaches focus on different moments in the data, the implications from both dimensions are not necessarily the same. This would suggest that the underlying economic model is missspecified and would motivate other model specifications. The estimation results in either dimension have implications for the process in the other dimension. It is interesting to compare the results and to test whether they are consistent with each other.

In LS (1992a) the underlying factor processes are estimated and the GMM procedure from Hansen (1982) is applied to test overidentifying cross sectional restrictions. The moments they include in the GMM procedure deal with changes in yields. Testing the model in first differences, however, requires only the estimation of four of the six parameters, since the parameters d_1 and d_2 are not identified. Their GMM test does not reject the overidentifying restriction and this suggests that the LS model adequately describes the yield curve dynamics. From a practical point of view, however, the approach is not desirable. Since not all parameters of the model are estimated, the term structure cannot be computed, which makes a comparison of yield curves with other non-nested model impossible. Moreover, the GMM approach heavily depends upon the moments that have been chosen. The fact that the changes in yields are adequately represented does not mean that the same holds for the yield curve itself. Finally, LS (1992a) state that the point estimates of the models estimated in their tests, may differ significantly from those obtained by procedures that

estimate all six parameters simultaneously, because of high correlation between the subset of parameters and the remaining parameters. In LS (1993a) they propose an estimation procedure based on the method of moments, in which they estimate all parameters.

Below, we describe and apply the time series oriented estimation procedure as proposed by LS (1993a). Also, the panel data approach as proposed by Bams and Schotman (1997) is set out. The approaches differ with respect to the type of data that is incorporated and also the particular moments on which the models are fitted are different.

5.3.1 Time series approach

LS (1993a) infer parameter values under the actual probability measure from the unconditional first and second moments of the underlying factors by applying the method of moments. The diffusion processes for the instantaneous spot rate and the volatility of the instantaneous spot rate in equations (5.5) and (5.6) determine the specification of the unconditional moments. For the instantaneous spot rate, r, the expected value and variance are given by

$$E(r) = -\left[\frac{d_1}{C_1} + \frac{d_2}{C_2}\right] \tag{5.23}$$

and

$$Var(r) = \frac{\beta_{11} d_1}{2C_1^2} + \frac{\beta_{22} d_2}{2C_2^2} \tag{5.24}$$

The unconditional moments for the associated volatility process are given by

$$E(V) = -\left[\frac{\beta_{11} d_1}{C_1} + \frac{\beta_{22} d_2}{C_2}\right] \tag{5.25}$$

and

$$Var(V) = \frac{\beta_{11}^3 d_1}{2C_1^2} + \frac{\beta_{22}^3 d_2}{2C_2^2} \tag{5.26}$$

Because the LS model consists of six unknown parameters, d_1, d_2, C_1, C_2, β_{11} and β_{22}, the method of moments requires six moments to identify them. The level effect relation between spot rate and variance in equation (5.7) leads to two more restrictions. LS (1993a) set the observed minimum value equal to the lowerbound, and the observed maximum value equal to the upperbound. This results in two additional moments that determine the volatility parameters

$$\beta_{11} = \min_t \left(\frac{V_t}{r_t}\right) \tag{5.27}$$

$$\beta_{22} = \max_t \left(\frac{V_t}{r_t}\right) \tag{5.28}$$

In the methods of moments estimation procedure, the population moments are replaced with the sample moments. To determine the sample moments of V_t, we first need to

estimate the time series processes of the underlying factors, given in equations (5.5) and (5.6). Following LS (1993a), we consider the discrete time representation of the continuous time processes, which is given by

$$r_t - r_{t-1} = a_0 + a_1 r_{t-1} + a_2 V_t + e_t \tag{5.29}$$

with

$$e_t \sim N(0, V_t) \tag{5.30}$$

where the variance, V_t, follows a GARCH process of the form

$$V_t = b_0 + b_1 e_{t-1}^2 + b_2 V_{t-1} + b_3 r_{t-1} \tag{5.31}$$

Since we apply monthly data to estimate the GARCH model, all units of time are monthly. In the specification a_0, a_1, a_2, b_1, b_2 and b_3 are unknown parameters that are estimated. This leads to an implied time series process for both the spot rate and for the volatility. The sample moments of these series are calculated and substituted for the moments in equations (5.23), (5.24), (5.25), (5.26), (5.27) and (5.28). The parameter estimates for the continuous time processes are implied by the six moments and follow almost immediately.

Recall that the estimates are for parameters under the actual probability measure, whereas for term structure implications and derivative securities pricing the risk adjusted parameters are needed. The proportionality constant in the market price of risk, λ, is required to infer the implied yield curve process from the time series parameter estimates. This market price of risk is only available from cross sectional data. We use the fact that the infinite yield, $Y(\infty)$, is of the form

$$Y(\infty) = \frac{d_1}{\beta_{11}}(\xi + C_1) + \frac{d_2}{\beta_{22}}(\psi + C_2 - \lambda) \tag{5.32}$$

In the empirical part we substitute the average value of the longest available yield for $Y(\infty)$ as an approximation. Together with the other parameter estimates this results in a value for the market price of risk.[5] Finally, this leads to an implied yield curve for each time, by substitution of the parameters in equations (5.8), (5.9), (5.10) and (5.11).

5.3.2 Panel data approach

In the panel data approach the yield curve itself is fitted directly, which results in direct estimates for the risk-adjusted parameters. The yield curve is observed for N different

[5] An alternative approach we have pursued to come up with a value for λ, is by performing a grid search over λ that results in an on-average best fit of the entire yield curve. This leads to similar results as given in the text.

maturities, denoted by $\tau_1 < \tau_2 < \cdots < \tau_N$. No assumptions or approximations need to be made for the unobserved factors, since they can be treated as unknown parameters which are estimated along with the structural parameters of the model. This is what we refer to as the panel data model with fixed effects. Other studies in which the unobserved factor have been treated as fixed effects are Brown and Dybvig (1986), Brown and Schaefer (1994), DeMunnik and Schotman (1994) and Bams and Schotman (1997). In the fixed effects case the factors are estimated only from yield curve information and no reference is made to the time series properties of the factors.

Although the LS model as given in equation (5.8) presumes an exact relationship between the yields and the factors, in practice this is not the case. Therefore, we specify an econometric model by the inclusion of an error term in the model.[6] Pooling the data for all yields leads to the following model specification

$$\begin{pmatrix} Y_t(\tau_1) \\ \vdots \\ Y_t(\tau_N) \end{pmatrix} = \begin{pmatrix} A(\tau_1) \\ \vdots \\ A(\tau_N) \end{pmatrix} + \begin{pmatrix} B_1(\tau_1) & B_2(\tau_1) \\ \vdots & \vdots \\ B_1(\tau_N) & B_2(\tau_N) \end{pmatrix} \begin{pmatrix} r_t \\ Y_t \end{pmatrix} + \begin{pmatrix} \varepsilon_t(\tau_1) \\ \vdots \\ \varepsilon_t(\tau_N) \end{pmatrix} \tag{5.33}$$

where $A(\tau)$, $B_1(\tau)$ and $B_2(\tau)$ depend in a specific way on the structural parameters as given in equations (5.9), (5.10) and (5.11).[7] In matrix notation the model reads

$$Y_t = A + Bf_t + \varepsilon_t \qquad t = 1, \ldots, T \tag{5.34}$$

where

$$A = \begin{pmatrix} A(\tau_1) \\ \vdots \\ A(\tau_N) \end{pmatrix} \tag{5.35}$$

and the factor loadings are summarized as

$$B = \begin{pmatrix} B_1(\tau_1) & B_2(\tau_1) \\ \vdots & \vdots \\ B_1(\tau_N) & B_2(\tau_N) \end{pmatrix} \tag{5.36}$$

The error term is given by

$$\varepsilon_t = \begin{pmatrix} \varepsilon_t(\tau_1) \\ \vdots \\ \varepsilon_t(\tau_N) \end{pmatrix} \tag{5.37}$$

[6] Other examples where error terms are included to an equilibrium specification are Jacquier and Jarrow (1995), Renault (1996), Brown and Dybvig (1986), Frachot, Lesne and Renault (1995), De Jong (1996) and others.

[7] For purposes of readability, we suppress the dependence in our notation.

Although the error term is not a part of the equilibrium model, it captures all forms of misspecification and measurement error. For the error term we assume that it is normally distributed and we adopt the covariance matrix that is proposed in Bams and Schotman (1997)

$$\varepsilon_t \sim N(0, \omega^2 S) \tag{5.38}$$

where the (i,j)-th element of S is of the form

$$S_{i,j} = \frac{\phi^{\lfloor \tau_i - \tau_j \rfloor}}{(\tau_i \tau_j)^d} \qquad i = 1, \ldots, N \qquad j = 1, \ldots, N \tag{5.39}$$

This specification takes account for the cross sectional heteroskedasticity in the data, since volatility is inversely related with time-to-maturity. It also takes account of the fact that yields that differ little in time-to-maturity show higher correlation, whereas yields that differ more in time-to-maturity are less correlated. It is a parsimonious specification, which allows the inclusion of many different yields.

Under the assumption of normally distributed error terms, the loglikelihood follows directly as

$$\ln L = -\frac{1}{2} NT \ln \omega^2 - \frac{1}{2} T \ln |S| - \frac{1}{2\omega^2} \sum_{t=1}^{T} (Y_t - A - B f_t)' S^{-1} (Y_t - A - B f_t) \tag{5.40}$$

The factors and the scale parameter in the covariance matrix are determined analytically, by taking the first order conditions. For ω^2 it follows that

$$\hat{\omega}^2 = \frac{1}{NT} \sum_{t=1}^{T} (Y_t - A - B \hat{f}_t)' S^{-1} (Y_t - A - B \hat{f}_t) \tag{5.41}$$

The factors are treated as fixed effects, which means that we treat each factor as an unknown parameter, that has to be estimated. The number of factors is usually large, since it is equal to the number of time points in our data set. A non-linear relation between yields and factors would lead to T additional incidental parameters, that cannot be substituted out and that have to be estimated along with the structural parameters by some non-linear optimization algorithm. However, because of the affine relation of the yields and the factors, the first order condition results in an analytical expression for the fixed effects estimator

$$\hat{f}_t = (B' S^{-1} B)^{-1} B' S^{-1} (Y_t - A) \qquad t = 1, \ldots, T \tag{5.42}$$

This allows us to substitute the factors out of the loglikelihood, which leads to a sizable reduction of the numbers of parameters. Substitution of equations (5.41) and (5.42) in equation (5.40) yields the concentrated loglikelihood

$$\ln L = -\frac{1}{2} NT \ln \hat{\omega}^2 - \frac{1}{2} T \ln |S| \tag{5.43}$$

where $\hat{\omega}^2$ is a non-linear function of the structural parameters and the market price of risk, and S is a non-linear function of the parameters in the covariance matrix specification, ϕ and d. These parameters are estimated by maximization of the loglikelihood, which is carried out by applying the Newton-Raphson optimization routine.

The conditional covariance matrix of the factors is given by

$$V_f = \hat{w}^2 \left(B' S^{-1} B \right)^{-1} \tag{5.44}$$

for each time period, $t = 1, \ldots, T$. For the remaining parameters the optimization routine provides a heteroskedasticity consistent covariance matrix, based on the gradient and the Hessian of the loglikelihood in the optimum.

5.4 Data

Interest rates are only directly observable for short maturities, the longer term yields are implicitly available in bond prices. The current market price of a bond reflects the discounted value of all future cash flows plus accrued interest. The rates at which all cash flows are discounted determines the yield curve.

The econometric analysis is greatly simplified if we can work with discount yield data instead of bond prices. In that way we avoid the non-linearities that stem from the fact that bonds normally carry coupon payments. In the literature a number of approaches have been pursued to construct yield curves. We have used the cubic spline method as proposed by McCulloch (1975) to construct yield curves.

The raw data set consists of US government bond data, available from the CRSP tapes for the period January 1970 until December 1994. The constructed yield curves are evaluated at 16 points with maturities ranging from 1 month to 6 months and 1 year to 10 years. This leads to a data panel which consists of $T = 300$ cross sections, where each cross section is represented by $N = 16$ discount yields with different maturities.

Figure 5.1 shows the entire data panel, which is constructed by evaluation of the fitted cubic spline functions in 16 points on the curve. Data on the short end are less reliable since they include all sorts of maturity effects and monetary policy effects. Summaries statistics for the yield levels are presented in table 5.1. The yields show a term structure that on average is increasing. The term structure of volatilities is decreasing, and from the last column we find that yields are highly autocorrelated. In table 5.2 another feature of yield curve data is illustrated. Yields that differ little in time-to-maturity are highly correlated. Moreover, the correlation decreases as the difference between the maturities of the yields gets larger.

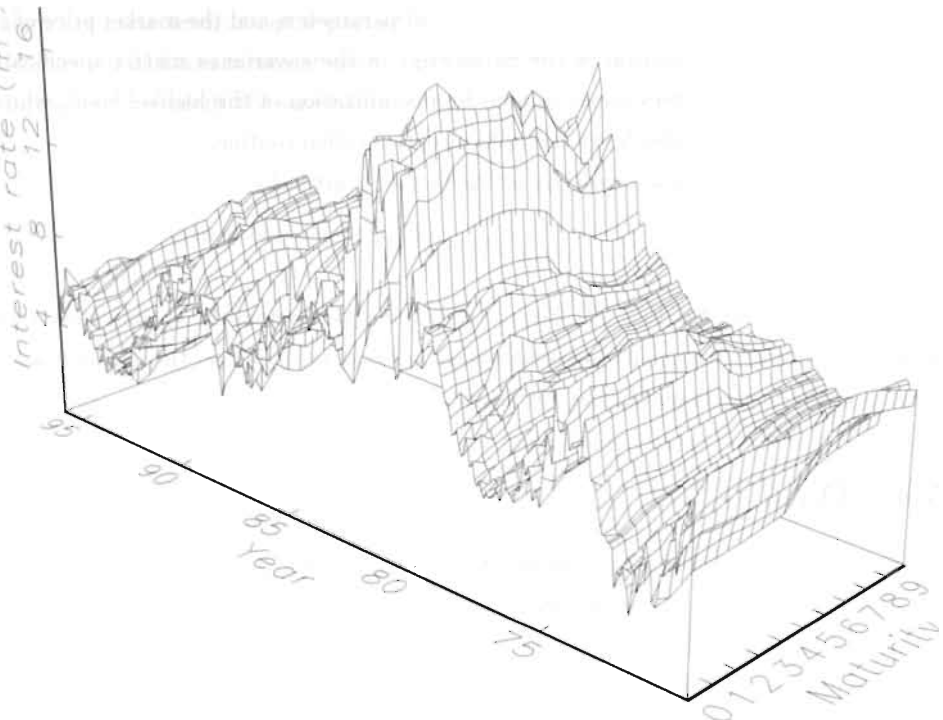

FIGURE 5.1: US INTEREST RATES

TABLE 5.1: SUMMARY STATISTICS

Maturity	Average	Standard Deviation	Minimum	Maximum	Auto Correlation
1 month	6.99	2.98	2.49	16.97	0.904
2 months	7.05	2.87	2.52	16.37	0.962
3 months	7.13	2.83	2.64	15.93	0.971
4 months	7.21	2.82	2.90	15.96	0.968
5 months	7.30	2.82	2.89	16.21	0.962
6 months	7.38	2.83	2.92	16.55	0.960
1 year	7.54	2.65	2.94	15.89	0.968
2 years	7.89	2.52	3.77	16.31	0.973
3 years	8.07	2.39	4.20	15.88	0.975
4 years	8.23	2.33	4.52	16.00	0.978
5 years	8.33	2.23	4.81	15.50	0.979
6 years	8.42	2.16	4.99	14.93	0.980
7 years	8.49	2.14	5.16	14.71	0.982
8 years	8.54	2.14	5.32	14.80	0.983
9 years	8.56	2.13	5.47	15.04	0.983
10 years	8.58	2.12	5.57	15.28	0.983

TABLE 5.2: CORRELATION MATRIX

τ	1/12	2/12	3/12	4/12	5/12	6/12	1	2	3	4	5	6	7	8	9	10
1/12	1.00	0.98	0.96	0.95	0.95	0.95	0.94	0.91	0.90	0.88	0.87	0.85	0.83	0.82	0.82	0.82
2/12	0.98	1.00	0.99	0.99	0.98	0.98	0.97	0.95	0.93	0.91	0.89	0.87	0.86	0.84	0.84	0.85
3/12	0.96	0.99	1.00	1.00	0.99	0.99	0.98	0.96	0.93	0.92	0.90	0.88	0.86	0.85	0.85	0.85
4/12	0.95	0.99	1.00	1.00	1.00	1.00	0.99	0.96	0.94	0.92	0.91	0.89	0.87	0.86	0.85	0.86
5/12	0.95	0.98	0.99	1.00	1.00	1.00	0.99	0.97	0.94	0.93	0.91	0.89	0.87	0.86	0.86	0.86
6/12	0.95	0.98	0.99	1.00	1.00	1.00	0.99	0.97	0.95	0.93	0.91	0.89	0.88	0.87	0.86	0.87
1	0.94	0.97	0.98	0.99	0.99	0.99	1.00	0.99	0.97	0.96	0.95	0.93	0.91	0.90	0.90	0.90
2	0.91	0.95	0.96	0.96	0.97	0.97	0.99	1.00	0.99	0.99	0.98	0.97	0.96	0.95	0.95	0.95
3	0.90	0.93	0.93	0.94	0.94	0.95	0.97	0.99	1.00	1.00	0.99	0.99	0.98	0.97	0.97	0.97
4	0.88	0.91	0.92	0.92	0.93	0.93	0.96	0.99	1.00	1.00	1.00	0.99	0.99	0.98	0.98	0.98
5	0.87	0.89	0.90	0.91	0.91	0.91	0.95	0.98	0.99	1.00	1.00	1.00	0.99	0.99	0.99	0.99
6	0.85	0.87	0.88	0.89	0.89	0.89	0.93	0.97	0.99	0.99	1.00	1.00	1.00	1.00	0.99	0.99
7	0.83	0.86	0.86	0.87	0.87	0.88	0.91	0.96	0.98	0.99	0.99	1.00	1.00	1.00	1.00	0.99
8	0.82	0.84	0.85	0.86	0.86	0.87	0.90	0.95	0.97	0.98	0.99	1.00	1.00	1.00	1.00	0.99
9	0.82	0.84	0.85	0.85	0.86	0.86	0.90	0.95	0.97	0.98	0.99	0.99	1.00	1.00	1.00	1.00
10	0.82	0.85	0.85	0.86	0.87	0.90	0.95	0.97	0.98	0.99	0.99	0.99	1.00	1.00	1.00	1.00

5.5 Results

In the empirical part three model approaches are compared with respect to their implications for the term structure of interest rates. The focus is on the fit of the yield curve. We consider three different approaches that differ either with respect to the number of factors underlying the term structure model or with respect to the moments of the data that are fitted in the estimation routine.

We consider the CIR one-factor model, which is estimated in a panel data framework. Comparison of this nested model with the LS panel data approach makes it possible to quantify the contribution of volatility as a second factor. We also estimate the LS model as proposed by LS (1993a). The focus in their approach is on the first and second unconditional moments of the underlying factors, which is different from the panel data approach. Comparison of this model with the LS panel data approach, shows how the implications for the yield curves are different when the LS model is fitted to other moments in the data. In the last part of this section, we compare the implications of the alternative models for the pricing of an interest rate cap.

To visualize the implications for the term structure, the average term structure and the term structure of unconditional volatilities, as implied by the models and as given by the observed yields, are depicted. The implied time series for the factors are given and in the case of the instantaneous spot rate it is compared with the observed one month spot rate. We provide summary statistics on the absolute values of the residuals for different yields, to quantify the accuracy of the estimated models. Finally, we provide values for the constant term and the factor loadings, to get an impression of the contribution of the factors.

5.5.1 CIR one-factor panel data model

In the CIR one-factor model, the underlying factor is the instantaneous spot rate. The model is a special case of the LS model and therefore serves as a benchmark. The focus is on the cross sectional information in the yield curve, which leads to a panel data model, in which the factors are treated as fixed time effects. The specification for the loglikelihood in the CIR model is a special case of the specification in equation (5.40), where A is replaced with A^* and B with B^*. The estimators for ω^2 and the factors follow after the same substitution in equations (5.41) and (5.42). In the estimation procedure the parameters d_1 and β_{11} are constrained to be positive and C_1 is constraint to be negative.

TABLE 5.3: ESTIMATION RESULTS CIR MODEL

Parameter	Estimate	Standard error
β_{11}	0.0124	0.0006
d_1	0.0053	0.0003
C_1	-0.0000	-
ϕ	0.3027	0.0037
d	0.2976	0.0054
lnL	7848	

Notes: The table reports the panel parameter estimates for the CIR one-factor model. The associated standard errors are heteroskedasticity and autocorrelation consistent. The parameter C_1 is set to its lowerbound of 1e-20, and hence its standard error cannot be computed.

The parameter estimates from the panel data model are given in table 5.3. The mean reversion parameter, C_1, is extremely low and the lowerbound of 1e-20, that we impose in the loglikelihood optimization routine, is binding. The reason for this low mean reversion is that the data contain a lot of parallel shifts and the different yield curves do not converge for the longest yields in the data set ($\tau_N = 10$). The variation in ten-year yields forces the mean reversion parameter to be extremely low. The model implies a term structure with an infinite yield of 6.8 percent.

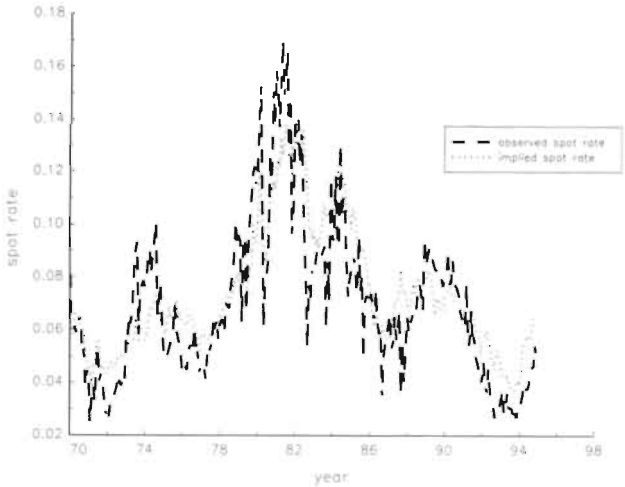

FIGURE 5.2: IMPLIED FACTORS

The figure shows the estimated time series parameters for the instantaneous spot rate for the CIR one-factor model, estimated in a panel data framework.

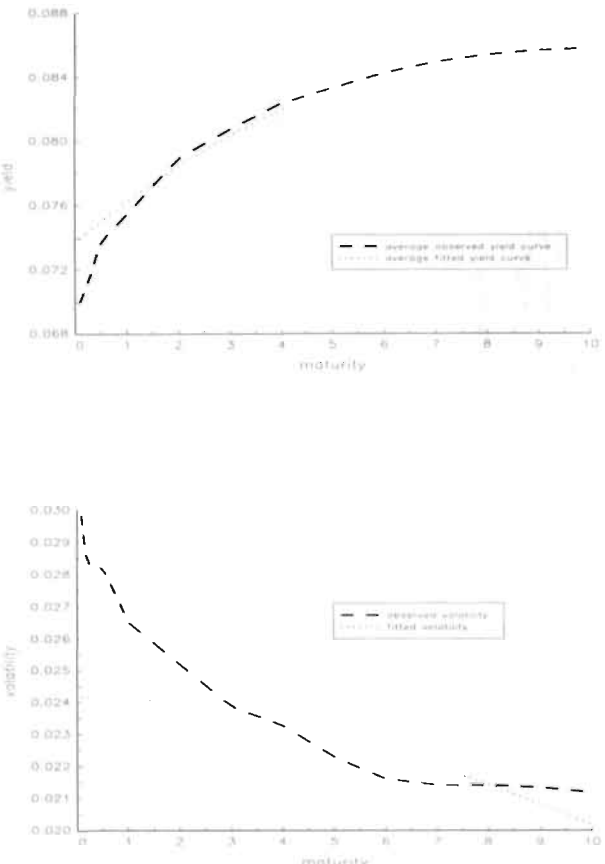

FIGURE 5.3: FIT TERM STRUCTURE

The top figure shows the observed and estimated unconditional average of the term structure of interest rates for the CIR single factor model. The bottom figure represents the unconditional term structure of volatilities.

The fixed effects estimates for the instantaneous spot rate are depicted in figure 5.2 together with the one month T-bill rate. The associated standard error equals 12 basis points for each fixed time effect. The estimated rate turns out to be much less volatile than the observed rate. A time series approach in which the T-bill rate is used as a proxy for the instantaneous spot rate, therefore is not consistent with the instantaneous spot rate as implied by the yield curve.

The fitted yield curves follow after substitution of the parameter estimates and the fixed effects estimates in equation (5.17). Figure 5.3 shows the fit of the average yield curve. At least for maturities longer than two years the observed and fitted average curves are close to each other, at the short end the average residual still is about 40 basis points. The term structure of volatilities at the bottom of figure 5.3 shows that especially for short term yields the fit of the unconditional volatility is extremely poor.

TABLE 5.4: SUMMARY STATISTICS RESIDUALS

Maturity	mean (in %)	autocorr
$\frac{1}{12}$	1.28	0.65
$\frac{2}{12}$	1.13	0.87
$\frac{3}{12}$	1.07	0.92
$\frac{4}{12}$	1.02	0.92
$\frac{5}{12}$	0.99	0.89
$\frac{6}{12}$	0.95	0.88
1	0.74	0.92
2	0.43	0.91
3	0.25	0.88
4	0.15	0.85
5	0.11	0.70
6	0.13	0.83
7	0.17	0.92
8	0.19	0.93
9	0.21	0.91
10	0.24	0.89

Notes: The table reports the average value of the absolute residuals and the autocorrelation of the residuals in the case we estimate the CIR one-factor model in a panel data framework, in which we treat the factor as fixed effects.

Table 5.4 gives summary statistics of the absolute values of the residuals for all maturities. Especially at the short end the absolute residual is extremely high, and it gets lower for longer term yields. The autocorrelation in the residuals is high for all maturities. This motivates the inclusion of an additional factor, to take account for the dynamics that are

present in the residuals.

TABLE 5.5: FACTOR LOADING

τ	$A^*(\tau)$	$B^*(\tau)$
$\frac{1}{12}$	0.00	1.00
$\frac{2}{12}$	0.00	1.00
$\frac{3}{12}$	0.00	1.00
$\frac{4}{12}$	0.00	1.00
$\frac{5}{12}$	0.00	1.00
$\frac{6}{12}$	0.00	1.00
1	0.00	1.00
2	0.01	0.99
3	0.01	0.98
4	0.01	0.97
5	0.01	0.95
6	0.02	0.93
7	0.02	0.91
8	0.02	0.89
9	0.02	0.86
10	0.02	0.83

Notes: The table reports the constants and the factor loading for the CIR one-factor model, estimated in a panel data model. All figures are in normal units.

In table 5.5 the constant and the factor loadings are given. At the short end the factor loading is close to unity, since the factor is the instantaneous spot rate. The factor loadings for longer term yields are lower, which implies that the behaviour of longer term yields is much more different from short term yields.

Overall, much more weight is put on the long end of the yield curve, which leads to a worse representation at the short end. Autocorrelation in the residuals already suggests the inclusion of a second factor. This second factor could take account for the variability at the long end of the yield curve and allow the instantaneous spot rate better to represent the short end of the yield curve. The results motivate the inclusion of the volatility factor in the LS model.

5.5.2 LS time series model

In this subsection we follow the method of moments approach, as proposed by LS (1993a). The algorithm focuses on the moments of the underlying factors. The instantaneous spot rate and the volatility of the instantaneous spot rate are not observable and hence have

to be approximated. The instantaneous spot rate is approximated with the one month yield and the volatility follows from the GARCH model in equations (5.29), (5.30) and (5.31). This specification is the discrete version of the continuous time factor processes in equations (5.5) and (5.6).

TABLE 5.6: ESTIMATION RESULTS LS MODEL, TIME SERIES APPROACH

Parameter	Estimate	Standard error
a_0	-0.0002	0.0002
a_1	-0.0060	0.0054
a_2	0.0006	0.0052
b_0	0.0000	-
b_1	0.4453	0.1245
b_2	0.2972	0.1079
b_3	0.0005	0.0001
e_0	0.0001	0.0225
r_0	0.0771	0.0084
h_0	0.0000	-
β_{11}	0.0007	-
β_{22}	0.0167	-
d_1	0.0002	-
C_1	-0.0273	-
d_2	0.0040	-
C_2	-0.4841	-
λ	-0.4656	-

Notes: The table reports the parameter estimates for the LS model with the one month spot rate to estimate the following GARCH model:
$r_t - r_{t-1} = a_0 + a_1 r_{t-1} + a_2 V_t + e_t$
$e_t \sim N(0, V_t)$
$V_t = b_0 + b_1 e_{t-1}^2 + b_2 V_{t-1} + b_3 r_{t-1}$

In table 5.6 the parameter estimates for this GARCH model are given. The initial residual term (e_0), interest rate (r_0) and volatility parameter (V_0) are not observed and are treated as unknown parameters. We impose restrictions that the unconditional average T-bill rate and the unconditional variance are positive, i.e. greater than or equal to a small positive number ($1e^{-6}$). Also the initial interest rate, and the initial volatility are imposed to be positive. Altogether, this leads to the following positivity constraints: $E(r) > 1e^{-6}$, $E(V) > 1e^{-6}$, $r_0 > 1e^{-6}$, $V_0 > 1e^{-6}$.

The parameter estimates for a_0, a_1 and a_2 deal with the change in interest level specification in equation (5.29). None of them differ statistically from zero which supports the random walk hypothesis for interest rates. The parameters b_0, b_1, b_2 and b_3 deal with the volatility specification of the interest rate. We find moderate persistence in volatilities (b_2)

and also the news effect (b_1) is significant. In figure 5.4 the time series process for the implied volatilities is depicted. The volatilities are annualized by multiplication with 12.

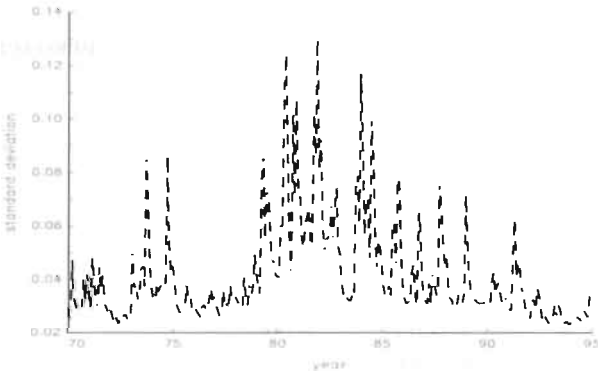

FIGURE 5.4: GARCH VOLATILITY

The figure shows the implied time series process for the standard deviation of the one month spot rate.

The method of moments, given by equations (5.23), (5.24), (5.25), (5.26), (5.27) and (5.28), is based on the implied time series of volatilities and the time series of the one month rate. We report the implied values for the structural parameters in table 5.6. To provide evidence on the implied yield curve, a value for the market price of risk is required. We infer λ from equation (5.32), where we approximate the infinite yield with the average value of the ten year yield.

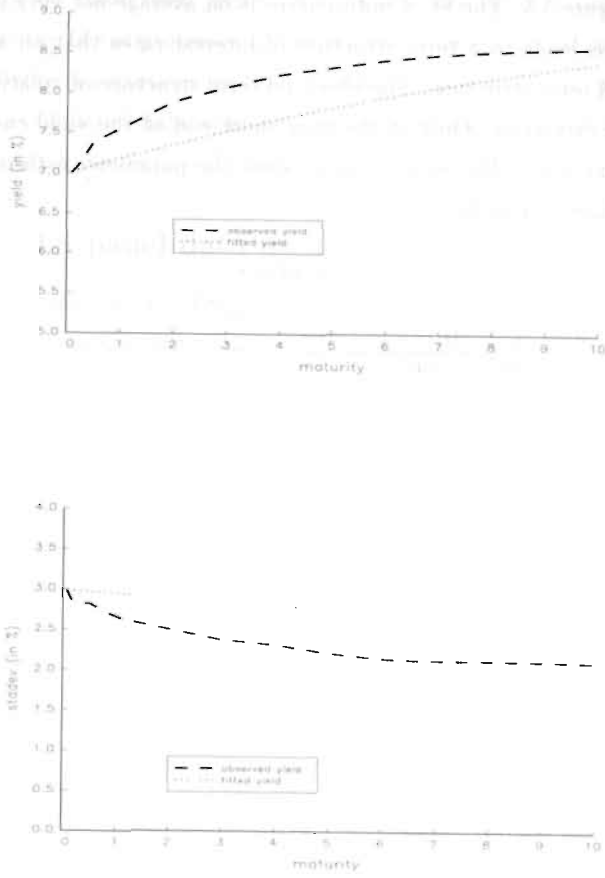

FIGURE 5.5: IMPLIED TERM STRUCTURE

The top figure shows the observed and implied average term structure of interest rates, the bottom figure shows observed and implied the term structure of unconditional variances, in case the factor processes of the LS model are estimated.

The implied average term structure of interest rates and the associated term structure of volatilities are given in figure 5.5. The fit of both curves is on average not very precise. The estimated factor process leads to a term structure of interest rates that on average underestimates the observed term structure. The observed term structure of volatilities is overestimated by the factor estimates. Only at the very short end of the yield curve the model matches the observed curves. This seems logical, since the parameter estimates are based on moments of the short rate only.

TABLE 5.7: FACTOR LOADINGS

τ	$A(\tau)$	$B_1(\tau)$	$B_2(\tau)$
$\frac{1}{12}$	0.00	1.00	0.02
$\frac{2}{12}$	0.00	1.00	0.04
$\frac{3}{12}$	0.00	1.00	0.06
$\frac{4}{12}$	0.00	1.00	0.07
$\frac{5}{12}$	0.00	0.99	0.08
$\frac{6}{12}$	0.00	0.99	0.09
1	0.00	0.99	0.11
2	0.01	0.97	-0.10
3	0.01	0.96	-0.59
4	0.01	0.95	-1.32
5	0.01	0.93	-2.24
6	0.02	0.92	-3.31
7	0.02	0.91	-4.47
8	0.02	0.90	-5.69
9	0.02	0.88	-6.94
10	0.02	0.87	-8.17

Notes: The table reports the constants and the factor loadings for the LS model in the case we estimate the model from the time series implied by the GARCH model. All figures are in normal units.

The constant term and factor loadings, implied by the estimates of the factor processes, are given in table 5.7. For yields with short time-to-maturity the factor loading of the spot rate almost equals one, and this factor loading decreases for longer maturities. The factor loading of volatilities is small but positive for short term maturities, and only for longer maturities the loadings are negative. From a theoretical point of view the relation between yields and spot rate volatility is assumed to be negative: investors are willing to pay more for securities that allow them to lock in a long-term guaranteed rate of return when the uncertainty about future money-market yields increases (LS, 1993b).

From this subsection we conclude that, when the parameters are estimated from the factor processes, the yield curves are only fitted well, on average, at the short end of the

curve. For longer term yields the fit is worse. This may stem from the fact that we only use data on short term yields, or from the way we determine the market price of risk. Because the factor processes are estimated under the actual probability measure, this does not provide us estimates for the risk adjusted probability measure. This motivates a cross sectional or panel data approach, which deals with parameters under the risk-adjusted probability measure.

5.5.3 LS panel data model

Again, we consider the LS model, but now we apply panel data techniques on the yields to infer the parameters under the risk adjusted probability measure. In table 5.8 the parameter estimates are reported.

TABLE 5.8: ESTIMATION RESULTS LS PANEL MODEL

Parameter	Estimate	Standard error
β_{11}	0.0000	-
β_{22}	0.0051	0.0015
d_1	0.0000	0.0010
C_1	0.0002	0.0049
d_2	0.0456	0.0171
C_2	-0.6147	0.0201
ϕ	0.6192	0.0209
d	0.5526	0.0063
lnL	10865	

Notes: The table reports the parameter estimates for the LS panel model, in which we treat the factors as fixed effects.

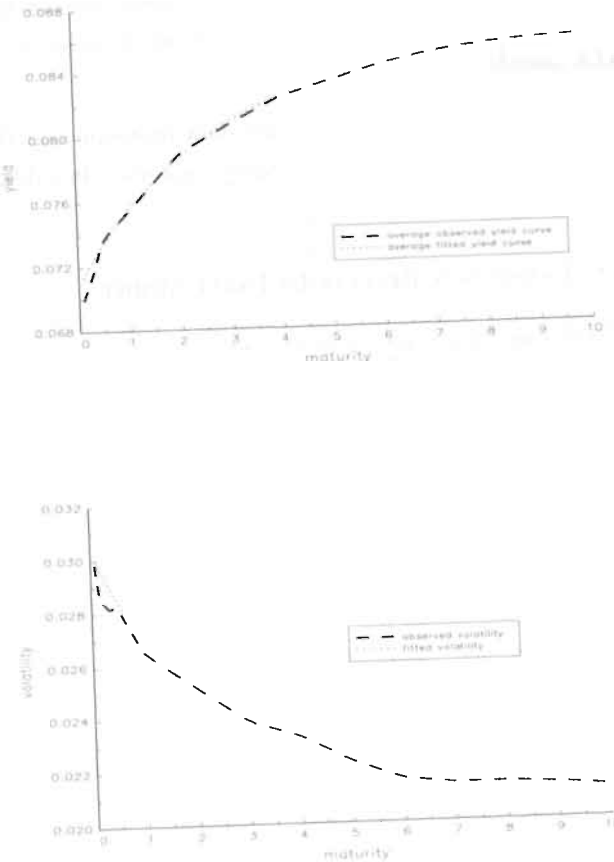

FIGURE 5.6: FIT TERM STRUCTURE

The top figure shows the observed and estimated unconditional average of the term structure of interest rates. The bottom figure represents the unconditional term structure of volatilities, for the LS panel model.

The LS model, estimated in a panel data framework, is capable of fitting the unconditional mean and volatility for all maturities on the yield curve, as is shown in figure 5.6. Only at the short end of the yield curve the fit is worse, but still it is a considerable improvement over the single factor CIR model.

TABLE 5.9: FACTOR LOADINGS

τ	$A(\tau)$	$B_1(\tau)$	$B_2(\tau)$
$\frac{1}{12}$	0.0019	1.00	-4.87
$\frac{2}{12}$	0.0037	1.00	-9.58
$\frac{3}{12}$	0.0054	1.00	-14.13
$\frac{4}{12}$	0.0071	1.00	-18.53
$\frac{5}{12}$	0.0087	1.00	-22.78
$\frac{6}{12}$	0.0103	1.00	-26.89
1	0.0188	1.00	-48.85
2	0.0314	0.99	-81.70
3	0.0402	0.99	-104.16
4	0.0465	0.98	-119.81
5	0.0510	0.98	-130.92
6	0.0544	0.97	-138.93
7	0.0569	0.96	-144.81
8	0.0589	0.96	-149.18
9	0.0605	0.95	-152.46
10	0.0618	0.95	-154.93

Notes: The table reports the constants and the factor loadings for the LS panel model, in which the factors are treated as fixed effects. All figures are in normal units.

The implied factor loadings and constant terms are given in table 5.9. We find that the spot rate is an important factor with almost a one-to-one relation between changes in spot rate and changes in yields. For longer term yields the one-to-one relationship weakens somewhat and the volatility of the spot rate becomes more important.

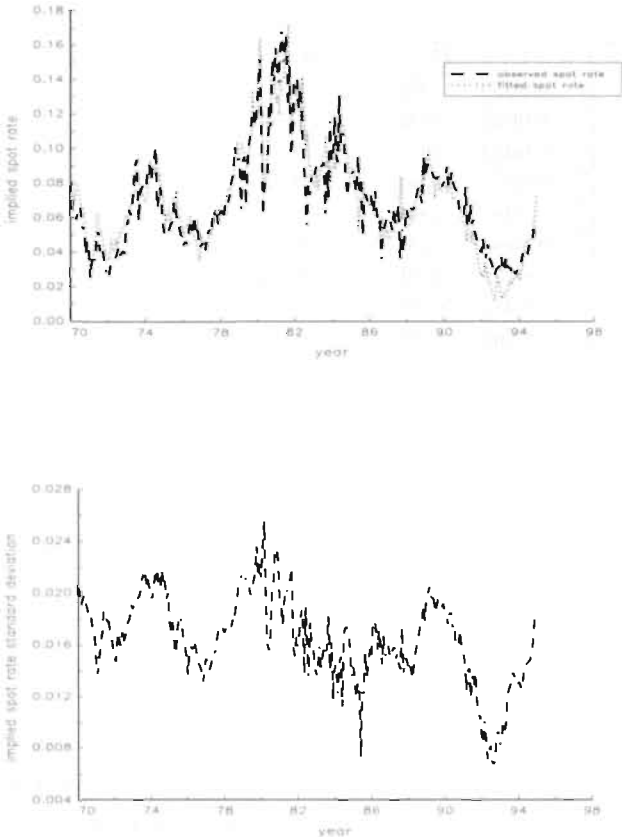

FIGURE 5.7: IMPLIED FACTORS

The top figure shows the estimated time series parameters for the first factor which is the instantaneous spot rate. In the bottom figure the standard deviation of the instantaneous spot rate is depicted, which is the square root of the second factor estimates, in case the Longstaff-Schwartz model is estimated from a time series approach.

The fixed effects estimates for both factors are given in figure 5.7, where the implied spot rate process closely resembles the actually observed one month rate. Also the associated standard errors, given by equation (5.44), for both the short rate and the volatility factors are low, 0.00363 and 0.0000242 respectively (where the average estimates for the factors are 0.0699 and 0.0004, respectively), which means that both are estimated quite precisely. The second factor allows the instantaneous factor more freedom to resemble the observed short rate. Since this second factor takes account for the behaviour at the intermediate and long end, the first factor is less determined by the long term maturities and better accounts for the short-term yield processes. Figure 5.7 shows that the estimated instantaneous spot rate closer resembles the observed one month rate, than in the CIR one-factor model as shown in figure 5.2.

TABLE 5.10: SUMMARY STATISTICS RESIDUALS

Maturity	Mean (in %)	Autocorr
$\frac{1}{12}$	0.93	0.19
$\frac{2}{12}$	0.65	0.38
$\frac{3}{12}$	0.50	0.59
$\frac{4}{12}$	0.40	0.70
$\frac{5}{12}$	0.37	0.59
$\frac{6}{12}$	0.34	0.45
1	0.19	0.24
2	0.11	0.27
3	0.12	0.46
4	0.10	0.57
5	0.11	0.49
6	0.11	0.72
7	0.11	0.85
8	0.08	0.87
9	0.04	0.72
10	0.08	0.84

Notes: The table reports the average value of the absolute residuals and autocorrelation of the residuals for the LS panel model, in which the factors are treated as fixed effects.

The residuals analysis in table 5.10 show that at the short end the average absolute residual is fairly large, whereas it decreases to low values as the time-to-maturity is longer. The same story holds for the associated standard deviation of the residuals. Finally, especially for longer term yields there is autocorrelation in the error terms. Loosely speaking, the second factor could be identified as a factor that is particular important for the curvature of the yield curve, whereas it allows more freedom to the spot rate factor to fit the

short end of the yield curve. The autocorrelation in the residuals at the long end of the yield curve suggests that a third factor is required to model the yield curve.

5.5.4 Cap pricing

So far, we argued that a time series dimension approach focuses on different moments in the data than a cross sectional approach. Also, the fit of the average yield curve from a time series point of view is worse in comparison with the fit from a cross sectional approach. Even though the panel data approach would be preferred for yield curve modeling, the estimates of a time series approach can still be interesting for pricing interest rate derivative securities, where volatility plays an important role. Typically volatility is well estimated from a time series point of view. In the cross sectional approach volatility is extracted from the convexity that is available in the yield curve, which a priori seems to lead to less fruitful results when the data would consist of many flat yield curves. In this subsection, we compare the implications for cap prices from the alternative estimation procedures.

In figure 5.8 the value of a cap is given as a function of the cap rate for the three alternative models. The current spot rate is set at $r_0 = 0.06$ and the current variance of the spot rate equals $V_0 = 0.0002$. The shapes of the curves are similar, only the levels are very different. Longstaff and Schwartz (1992b) find the same shape and explain the quick drop-off as the cap rate increases from the mean reversion in the interest rate model, because mean reversion implies that the probability of a cap with a high cap rate being at the money is much less than it would be in the absence of mean reversion.

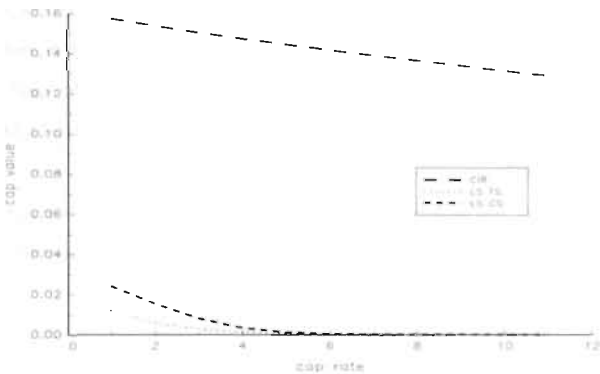

FIGURE 5.8: CAP RATES

The figure shows the relation between the cap rate and the value of the cap for the CIR model, for the LS time series approach and for the LS panel data approach. The intitial values for the spot rate and the variance of the spot rate are 0.06 and 0.0002, respectively.

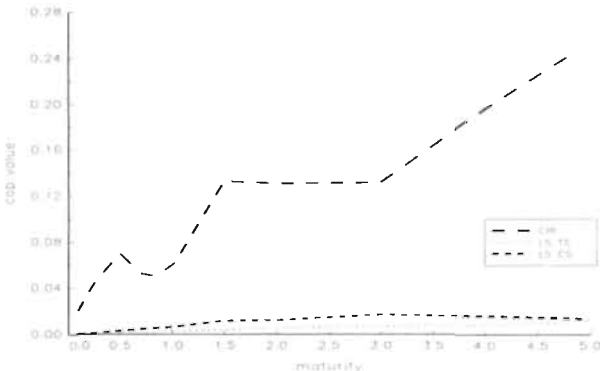

FIGURE 5.9: CAP MATURITIES

The figure shows the relation between the time to maturity of the cap and its value for the CIR model, for the LS time series approach and for the LS panel data approach. The intitial values for the spot rate and the variance of the spot rate are 0.06 and 0.0002, respectively.

In figure 5.9 the value of the cap is given as a function of the maturity of the cap, where the cap rate is held constant at 7 percent. Longstaff and Schwartz (1992b) explain the hump-shaped pattern of time-decay in the value of the cap because the variance of future values of the spot rate does not grow linearly with time. Instead it grows at a slower than linear rate because the model implies that the spot rate has a steady state distribution. This effect, in conjunction with the higher discount factor as time-to-maturity increases, results in the humped shape pattern.

Although for the three different parameter cases the shape of the cap value are similar, the actual values are very different and actual data on caps is required to assess which of the three cases is preferred.

5.6 Concluding remarks

We have considered the parameter estimates of the LS model from two different sources: time series data of the factor processes and cross sectional yield curve data. The time series approach provides a bad fit of the yield curve, except at the very short end. When we calibrate the model on yield curves directly, the fit is worse at the very short end. In comparison with the CIR one-factor model, we find that inclusion of spot rate volatility as a second factor contributes to the long part of the yield curve and therefore allows the spot rate to represent the short end of the yield curve better.

The LS model incorporates the same factors as in Andersen and Lund (1997). Same as in Andersen and Lund (1997), the short end of the yield curve is well presented by the LS term structure model, when the model is estimated using data on the one month rate only. The results from the panel data model show however that the LS two factor model is not able to represent the yield both at the short end and on the long end, simultaneously. This means that the model is misspecified and it motivates a model extension.

A second issue concerns the information that is included in estimating the LS model. Estimating term structure models from both dimensions simultaneously would of course give a more complete treatment of the problem than considering only a time series or panel data approach, but it is computationally cumbersome.[8] Andersen and Lund (1997), who estimated very general time series processes of the spot rate conclude that it would indeed be desirable to conduct a more formal statistical analysis of the theoretical versus

[8]Bams and Schotman (1997) estimate the Vasicek model in a panel data model, where the spot rate process is included as random effects. Because of the simple specification of the Vasicek spot rate process an analytical expression for the likelihood is available. They found that the parameter estimates from both dimensions are not consistent with each other and that cross sectional estimates are more precisely.

observed yield curves, but since in their framework evaluation of bond prices can only be done with numerical methods, it is a computationally challenging task. In our approach the cross sectional implications from the model are easily obtained. Taking account of the time series dimension would be possible by treating the factors as random effects which have non central chi-square distributions and use techniques. However, since the estimates from both dimensions are very different, this will not alleviate the misspecification of the model. The empirical results imply that the dynamics at the short end are very different from the dynamics at the long end.

Another possible extension of the LS model deals with the autocorrelation which is present in the long term residuals. The residual analysis indicates that possibly a third factor is present in yield curve data. The factor analysis in Litterman and Scheinkman (1991) has already shown that with three factors the yield curve is adequately described. An improvement over their approach would be to derive a three factor model within the equilibrium framework of CIR (1985b) and LS (1992a). The equilibrium approach makes the model both interpretable and applicable for the pricing of other interest rate derivative securities. Within the context of the LS model, it is interesting to include long term yield volatility as a third factor. This would allow the second factor, to take account for the fluctuations at the short end of the yield curve, whereas the third factor would deal with the fluctuations at the long end of the yield curve.

Chapter 6

Summary and Concluding Remarks

6.1 Introduction

In this thesis the term structure of interest rate has been investigated. The literature on the term structure of interest rate is huge and deals with many interesting issues. An important part of the literature has dealt with the theoretical development of models that consistently represent the behaviour of different yields simultaneously, that are applicable for the pricing of interest rate sensitive products or for forecasting purposes. Another part of the literature has dealt with the empirical properties and implications of yield curve models. Important issues in this field are the methods that are proposed to estimate term structure models and a comparison of the implications from different models or methods.

We have been concerned with the affine class of term structure models, which is interesting for its theoretical properties, since it describes the cross sectional relations of all yields by taking account of no-arbitrage restrictions. At the same time, we recognize that exact relationships never hold in practice and hence we propose to extend this type of models by explicitly incorporating an error term. The focus has been on the empirical properties of the resulting econometric model. Important subquestions of an empirical analysis deal with a motivation for the properties of the error term, which moments of the data are used to estimate the model, a comparison of the in-sample behaviour of alternative models and implications for other interest rate sensitive securities (which could be considered as an out-of-sample comparison of the models).

The choice of the affine class of term structure models leads to an attractive framework for the econometric analysis. The cross sectional relations between yields are available in closed form expressions and the model is linear in the underlying factors. This representation visualizes that term structure models have two dimensions: a time series dimension, which is represented by the underlying factor dynamics, and a cross sectional dimension,

given by the relation between different yields at a particular point of time. We propose to employ a panel data analysis focusing on the cross sectional information of all yields for all times simultaneously. This leads to efficient estimates of the parameters of interest (i.e. the parameters under the risk adjusted probability measure). At the same time the panel data approach is easily comparable with a time series approach of the underlying factor, since there exists a one-to-one relationship between the two dimensions, after an assumption about the market price of risk.

In the next sections we summarize the main findings of this thesis and we end this chapter with some directions for further research.

6.2 Summary

This thesis proposes to use an econometric version of the affine class of term structure models and furthermore it advocates to employ panel data techniques to estimate this class of models. The empirical properties are investigated for US Government bonds. The empirical findings are promising since they lead to efficient and plausible estimates for the parameters of such models. The framework allows for (1) an in-sample comparison of the various models within this class, (2) a comparison of the panel data approach with time series approaches, (3) testing of model restictions and assumptions on the market price of risk, (4) out-of-sample implications of other interest rate derivative securities.

In chapter one the theoretical framework for the affine class of term structure models is set out. Basically, it has been shown that an affine representation for the dynamics of some underlying factors, leads, under no-arbitrage conditions or within an equilibrium framework, to a model for discount yield curves that is closed form and affine in the underlying factors. An explicit assumption about a market price of risk has been incorporated to link the factor processes, which are expressed in terms of parameters under the actual probability measure, to the yield curve model, which is expressed in terms of parameters under the risk adjusted probability measure. This class of models is general, in a sense that it includes both endogenous and exogenous models in the same framework.

The assumptions underlying the derivation of the affine class of term structure models already lead to many interesting questions. For example, is it reasonable to assume that there exists an exact relationship among all yields and between the underlying factor dynamics and the associated yield curve? Are the asssumptions for the market price of risk plausible? Are the data consistent with the cross sectional restrictions that are implied by the model? And, is the underlying factor process also applicable for the pricing other interest rate derivative securities?

In chapter two different methods are compared to construct discount yield curves from coupon bearing bonds. Because usually term structure models are expressed in terms of discount yields, the analysis of these models is greatly simplified if we have the disposal of discount yields. A direct analysis of coupon bearing bonds hampers the econometric analysis since it leads to non-linear relationships between bond prices and discount yields. Moreover, the model implications are less transparant: residuals on coupon bearing bonds are less informative about the source of the misfit than in the case of discount yields. The usual approach is to transform the information that is available in coupon bearing bonds in terms of discount yields, using some parametric specification for the implied yield curve at each cross section. This results in a regular panel of discount yields that serves as input for the empirical analysis in the subsequent chapters.

We compare the Cubic Spline method with the Nelson-Siegel method to parametrize the discount yield curve and we compare spread weighting and duration weighting to weight the error terms in the fit criterion. General findings are that the Nelson-Siegel approach leads to smoother yield and forward curves than the Cubic Spline method. The spread weighted yield curves showed a lot of outliers, even after the worst performing bonds were filtered out. On the basis of these findings the preferred method would be the duration weighted Nelson-Siegel method. Care should however been taken that the method does not smooth away too much information present in the coupon bearing bonds. In the subsequent analysis the empirical results have always been compared for differently constructed yields. Another important finding is that all method results in less reliable implied short term yields. The confidence intervals around the implied yield curves are wide at the short end of the yield curve, possibly due to monetary policy effects or maturity effects. Also for maturities longer than ten years the yields are unreliable, due to the limited number of observations we have for long term bonds.

In chapter three the information that is available in a panel of yield curves is investigated without the assumption of an explicit term structure model. Apparently a panel of yields is more informative than an univariate time series of some yield. We revisit the problem of testing the expectations hypothesis (EH) of the term structure of interest rates. In the literature rejection of the EH has been attributed to small sample bias and/or to time varying risk premia. Instead of using univariate time series, in a panel data approach all time series of yields are included simultaneously to test the EH. An advantage of the panel data approach, hence, is efficiency gain. However, also in a panel data framework the expectations hypothesis is rejected.

Moreover, the panel data approach facilitates the explicit inclusion of a model for the risk premium. Since the risk premium is not observable, the usual approach is to model it

either as an unknown parameter, which leads to a fixed effect estimator, or as a random variable, which leads to a random effect estimator. We model the cross sectional relation in the risk premium with a fixed effects estimator. The results show that the premium increases with maturity. The time series behavior of the risk premium is modeled as an AR(1) process under the assumption that the risk premium is a common factor in yields of all maturities. We find a considerable degree of persistence in the risk premium. Often term structure models are derived under the assumption of a constant market price of risk. The analysis has shown that such an assumption is in conflict with the empirical findings in chapter three

In chapter four we set up an econometric analysis of affine term structure models. We propose a panel data analysis of the yield curve, where only cross sectional information is incorporated and no reference is made to the dynamics of the underlying factor process. No assumption about a market price of risk has to be made and this leads directly to parameter estimates under the risk adjusted probability measure. For pricing purposes only the parameters under the risk adjusted probability measure are of interest.

Since in practice exact relations between yields will not hold, an error term is included that takes account for model and measurement error. To keep the model tractable we parameterize the covariance matrix of the error term. This parameterization takes account of the inverse relation between volatility and time to maturity, and it models the cross sectional relation between residuals with different time to maturity. Finally, the error term specification accounts for autocorrelation in the error terms.

We consider the Gaussian term structure model, for which one and two factor endogenous and exogenous term structure models are investigated in depth. The empirical findings are that the panel data approach lead to very precise estimates of the parameters. In particular, mean reversion is estimated very precisely in the panel data approach in comparison with time series estimates. In comparison with approaches in which models are estimated from one cross section, we no longer find erratic and non-sensical parameter estimates when we employ a panel data approach. The in-sample fit of all models is precise for longer maturities, both for the average yield curve and for the term structure of unconditional volatilities. A second factor is required to take account for the dynamics of short term yields. Another attractive feature of the panel data approach is that the fixed effects estimates for the factors imply a time series behaviour of the instantaneous spot rate that resembles the behaviour of the one month T-bill rate. So, even though no time series information has been included the cross sections are informative on the time series dynamics.

In chapter five the Longstaff-Schwartz term structure model has been considered. The

model is promising both from a time series perspective and from a panel data perspective. Previous research has shown that the specifications of the two underlying factors turned out to provide adequate descriptions of the factor dynamics. In chapter four we found that a related two factor Gaussian term structure provides an adequate description of the yield curve. Because both approaches focus on different moments in the data it is interesting to compare the model implications from both approaches. We also compare the results with the CIR one factor model, which is nested in the Longstaff-Schwartz model.

The time series approach proves to give an adequate description of the first and second moment at the short end of the yield curve, but overall it underestimates the average yield curve and overestimates the term structure of volatilities. The panel data approach provides an adequate fit of the term structure except at the very short end. Although the two approaches show different empirical results, the question remains if also the implications for other derivative securities are different. By means of example we compared the models in their ability to price interest rate caps. It turns out that both models react in the same way to interest rate changes and to changes in time to maturity. However, cap prices implied by the time series approach are systematically higher than cap prices implied by a panel data approach. The price differences are relatively small, the differences with the CIR one factor cap prices are much higher. Actual observation of cap prices are required to give a conclusive answer about the preferred approach.

6.3 Directions for further research

In chapter three, four and five the results indicate that the results for short term yields, until maturities of six months, are most of the time different from the results for long term yield. In the literature this difference is often explained by the fact that there exists a lot of transitory noise that arises because of monetary policy actions with short term interest rates. From chapter two, however, we know that the short term yields are estimated with large standard errors and hence are not very precise. In most empirical work, constructed yield curves are treated as if the point estimates are exactly observed data. We propose to explicitly include the covariance matrix of the yields as implied by the curve fitting method. This suggests an approach in which the constructed yields still are treated as observed data, and in which the covariance matrix of these constructed yields is incorporated as the covariance matrix of the error term. This approach explicitly controls for the unreliability of the data at the short end of the yield curve. Issues that become important are (i) if rejection of the expectations hypothesis is mainly due to noise in the data, (ii) whether a seconds factor still is required to model the short end of the yield curve, and (iii) whether

time series oriented approaches really lead to different results than panel data approaches.

In chapter three we have adopted a model specification for the risk premium. This risk premium is closely related to the market price of risk, which is the expected instantaneous return on a discount bond in excess of the risk free rate, divided by the volatility of the instantaneous bond return. The discrete version of the instantaneous return on a discount bond is the holding period return as defined and modeled in chapter three. Extending the model with a time varying volatility specification for the holding period return may lead to inferences on the market price of risk. The market price of risk is an important ingredient in deriving models for derivative securities. Its properties are crucial for term structure modeling, since normally strong assumptions are made about its behavior.

Panel data estimates of affine term structure models are attractive because of the precise estimates that we obtain for parameters under the risk adjusted probability measure. For derivatives pricing the mean reversion parameter and the volatility parameter are particularly important. We found that a panel data approach is much more informative on the level of mean reversion than is usually found from a purely time series oriented approach. For the volatility, however, time series estimates are usually far more informative. In a panel data approach only the curvature of the term structure is informative on the level of volatility. Because we pool many different yield curves, we still find sensible values for the volatility parameter. It is interesting to investigate if both approaches can be combined to obtain efficient estimates for both parameters. However, a direct approach in which the time series process of the underlying factors is explicitly incorporated leads to the undesirable circumstance that an assumption about a market price of risk has to be made.

In this dissertation the models are evaluated on the in-sample fit of the average yield curve, the term structure of volatilities and the properties of the residuals. We found differences between one and two factor models, exogenous and endogenous model, and time series and cross sectional approaches. It remains however the question how severe these differences are for the pricing of other derivative securities. An interesting way of testing the model out-of-sample is by considering the implications of these model estimates for interest rate derivative pricing. First, it is interesting to find out whether the implied prices differ a lot or whether it does not matter too much which approach is pursued. Comparison of these implied derivative prices with actually observed prices, provides information on which parameters are crucial for a good fit. A panel data approach seems promising, since it does not require an assumption about a market price of risk and since it leads to efficient estimates of the relevant parameters.

Nederlandse samenvatting / Dutch summary

Dit proefschrift behandelt een econometrische versie van de affine klasse van termijnstructuur modellen. Panel data technieken worden gebruikt om deze klasse van modellen te schatten. We onderzoeken de empirische eigenschappen voor Amerikaanse staatsobligaties. De empirische resultaten zijn bemoedigend omdat panel data technieken leiden tot efficiënte en plausibele schatters voor de parameters van dergelijke modellen. In dit econometrisch raamwerk is het mogelijk om (1) een in-sample vergelijking uit te voeren van verschillende modellen in deze klasse, (2) een vergelijking te maken tussen de panel data aanpak en tijdreeksmodellen, (3) om model restricties en veronderstellingen omtrent de marktprijs van risico te toetsen, en (4) om de out-of-sample implicaties voor andere rente derivaten te onderzoeken.

In hoofdstuk 1 wordt het theoretische kader voor de affine klasse van termijnstructuur modellen uiteengezet. Er wordt aangetoond dat een affine representatie voor een onderliggend factor proces, onder de aanname van no-arbitrage condities of in een evenwichtsmodel, leidt tot een model voor de rentecurve dat een gesloten vorm heeft en dat affien is in de onderliggende factoren. Een expliciete veronderstelling over de marktprijs van risico is opgenomen om het factorproces, dat uitgedrukt wordt in termen van parameters onder de werkelijke kansmaat, te koppelen aan een rente termijnstructuur model, dat uitgedrukt wordt in termen van parameters onder de voor risico aangepaste kansmaat. De affine klasse bestaat zowel uit modellen die exogeen zijn als endogeen.

De veronderstellingen die ten grondslag liggen aan de afleiding van de affine klasse van termijnstructuur modellen, roept reeds een aantal interessante vragen op. Is het bijvoorbeeld aannemelijk om te veronderstellen dat er een exacte relatie bestaat tussen rentes met verschillende looptijden, en tussen een onderliggende factor dynamiek en de geassociëerde rente termijnstructuur? Zijn de veronderstellingen over de marktprijs van risico plausibel? Zijn de waargenomen rentes consistent met de cross-sectionele restricties die worden opgelegd door het model? En, is het onderliggende factor proces ook toepasbaar voor het

prijzen van andere rente derivaten?

In hoofdstuk 2 worden verschillende methoden vergeleken om een verdisconterings rente curve te construeren uit coupondragende obligaties. Omdat termijnstructuur modellen gewoonlijk worden uitgedrukt in termen van verdisconterings rentes, zal de analyse van deze modellen sterk vereenvoudigd worden als we daadwerkelijk de beschikking hebben over verdisconterings rentes. Een directe analyse van coupon dragende obligaties bemoeilijkt de econometrische analyse omdat het tot niet-lineaire relaties leidt tussen obligatieprijzen en verdisconterings rentes. Verder zijn in dat geval de model implicaties minder transparant: residuen van coupondragende obligatie prijzen zijn minder informatief over de aard van een misfit, dan in het geval van verdisconterings rentes. De gebruikelijke aanpak is om de informatie die aanwezig is in coupondragende obligaties uit te drukken in termen van verdisconterings rentes, door gebruik te maken van een of andere specificatie voor de geimpliceerde verdisconterings rentecurve voor iedere cross-sectie. Dit resulteert uiteindelijk in een panel van verdisconterings rentes dat als input dient voor de empirische studies in de volgende hoofdstukken.

We vergelijken de Cubic Spline methode met de Nelson-Siegel methode om de rente curve te parameterizeren en we vergelijken de spread tussen bied en laat koersen van de obligaties en de duration om de storingstermen te wegen in het fit criterium. De resultaten tonen dat de Nelson-Siegel methode tot gladdere curves leidt zowel voor de rente curve als voor de forward curve, dan in het geval van de Cubic Spline methode. De resultaten in het geval van spread weging leidt tot rente curves met veel meer uitschieters, zelfs nadat de slechts presterende obligaties uit de data set zijn weggefilterd. Deze bevindingen leiden ertoe dat de duration gewogen Nelson-Siegel methode de geprefereerde methode is. Voorzichtigheid is geboden dat de methode niet te veel informatie wegfiltert die aanwezig is in coupondragende obligaties. In alle volgende analyses zijn de resultaten altijd vergeleken voor rentecurves die op de verschillende manieren zijn geconstrueerd. Een ander belangrijk resultaat is dat alle methoden resulteren in minder betrouwbare geimpliceerde kortlopende rentes. De betrouwbaarheidsintervallen rondom de geconstrueerde rentecurves zijn breed aan de korte kant van de curve, mogelijk vanwege monetaire politiek of door allerlei effecten vlak voor het aflopen van obligaties. Ook rentes met looptijden langer dan tien jaar zijn onbetrouwbaar, vanwege het gelimiteerde aantal waarnemingen voor langlopende obligaties.

In hoofdstuk 3 is onderzocht welke informatie er aanwezig is in een panel van rente termijnstructuren, zonder dat er een expliciet model wordt verondersteld. Een panel van termijnstructuren bevat meer informatie dan een univariate tijdreeks van een of andere rente. We toetsen de Verwachtingswaarde Hypothese voor de rente termijnstructuur. In de

literatuur is verwerping van de Verwachtingswaarde Hypothese onder andere toegeschreven aan onzuiverheid in kleine steekproeven en/of aan tijdsvariërende risicopremies. Anders dan bij een univariate tijdreeks, wordt in een panel data aanpak alle rente tijdreeksen simultaan gebruikt om de Verwachtingswaarde Hypothese te toetsen. Een voordeel van de panel data aanpak is om die reden een toename in efficiëntie. Echter ook bij een panel data aanpak wordt de Verwachtingswaarde Hypothese verworpen.

Verder stelt de panel data methode ons in staat om de risico premie expliciet te modelleren. Omdat de risicopremie niet waarneembaar is, is de gebruikelijke aanpak om haar als een onbekende parameter te modelleren, wat resulteert in een fixed effect schatter, of als een kansvariabele, wat leidt tot een random effect schatter. De cross-sectionele relatie in de risico premie wordt als een fixed effect gemodelleerd. De resultaten laten zien dat de premie toeneemt met looptijd. Het tijdreeksgedrag van de risico premie wordt gemo-delleerd als een AR(1) proces onder de veronderstelling dat de risico premie een gezamelijke factor is in rentes met verschillende looptijden. De resultaten laten zien dat risico premies tijdsvariërend en behoorlijk persistent zijn. Vaak worden termijnstructuur modellen afgeleid onder de veronderstelling van een constante marktprijs voor risico. De analyse heeft laten zien dat zo een veronderstelling in strijd is met de empirische bevindingen in hoofdstuk 3.

In hoofdstuk 4 is een econometrische analyse uiteengezet voor de affine klasse van termijnstructuur modellen. Een panel data aanpak wordt gemotiveerd, waarin alleen cross sectionele informatie wordt gebruikt en waarin niet gerefereerd wordt naar de dynamiek van de onderliggende factoren. Er hoeft geen veronderstelling te worden gemaakt voor de marktprijs van risico en dit leidt direct tot parameter schattingen onder de voor risico aangepaste kansmaat. Voor het prijzen van rente producten zijn alleen de parameters onder de voor risco aangepaste kansmaat van belang.

Omdat in de praktijk exacte relaties tussen rentes met verschillende looptijden niet gelden is een storingsterm geintroduceerd die rekening houdt met model- en meetfouten. Om het model handelbaar te houden hebben we de covariantie matrix van de storingsterm geparametrizeerd. Deze parameterizatie houdt rekening met de inverse relatie tussen volatiliteit en looptijd, en ze modelleert de cross-sectionele relatie tussen residuen met verschillende loopduur. De storingsterm specificatie houdt ook rekening met mogelijke autocorrelatie in de storingstermen.

We voeren een uitgebreide analyse uit voor het Gaussiaans termijnstructuur model, waarvoor we één en twee factor endogene en exogene modellen schatten. De empirische bevindingen zijn dat de panel data methode tot precieze schattingen voor de parameters leidt. Met name de parameter die het mean reversion gedrag beschrijft wordt heel precies bepaald in een panel data aanpak in vergelijking met tijdreeks schattingen. In vergelijking

met methoden waarin modellen worden geschat per cross-sectie, vinden we in de panel data aanpak niet langer absurde parameter schattingen. De in-sample fit in alle modellen is precies voor rentes met langere looptijden, zowel voor de gemiddelde rentecurve als voor de termijnstructuur van onconditionele volatiliteiten. Een tweede factor is nodig om ook de dynamiek van rentes met kortere looptijden goed te modelleren. Een andere aantrekkelijke eigenschap van de panel data aanpak is dat de fixed effect schattingen voor de onderliggende factoren een tijdreeks proces voor de instantane rente impliceren die sterk overeenkomt met het gedrag van de waargenome één maands rente. Dus, alhoewel geen tijdreeksinformatie is opgenomen is de cross-sectie informatief over de tijdreeks dynamiek van de korte rente.

In hoofdstuk 5 wordt het Longstaff-Schwartz termijnstructuur model onder de loep genomen. Het model is veelbelovend zowel van een tijdreeks perspectief als van een panel data perspectief. Eerder onderzoek heeft aangetoond dat de specificatie van de twee onderliggende factoren in dit model een goede beschrijving geeft van de dynamiek van de factoren. In hoofdstuk 4 hebben we aangetoond dat het gerelateerde Gaussiaanse twee factor model een bevredigende beschrijving geeft van de rente curve. Omdat beide methoden zich concentreren op andere momenten in de data, is het interessant om de model implicaties van beide methoden te vergelijken. We vergelijken het resultaat tevens met het CIR één factor model, dat een speciaal geval is van het Longstaff-Schwartz model.

De tijdreeks methode resulteert in een adequate beschrijving van het eerste en tweede moment aan de korte kant van de rentecurve, maar in het algemeen onderschat het de gemiddelde rente curve en overschat het de termijnstructuur van onconditionele volatiliteiten. De panel data aanpak geeft een adequate beschrijving van de gehele rente termijnstructuur, behalve voor zeer korte looptijden. Alhoewel de verschillende methoden tot verschillende empirische resultaten leiden, blijft het de vraag of de implicatie voor andere rente derivaten ook sterk verschillen. Bij wijze van voorbeeld hebben we vergeleken hoe de modellen rente caps prijzen. Het blijkt dat in beide gevallen de prijzen van caps op dezelfde manier reageren op rente veranderingen en op veranderingen in de looptijd. Cap prijzen die bepaald worden in de tijdreeks methoden zijn systematisch hoger dan cap prijzen die geimpliceerd worden in een panel data aanpak. De verschillen in de prijzen zijn relatief klein, de verschillen met de prijzen die volgen voor het CIR één factor model zijn veel hoger. Daadwerkelijk waargenomen cap prijzen zijn nodig om een uitspraak te doen aangaande welke methode verkozen dient te worden.

Bibliography

ADDOLORATO, F. AND A. BERARDI (1994), Long-term Volatility and the Term Structure: Evidence from Italian Treasury Bonds, *Workingpaper*.

ANDERSEN, T.G. AND J. LUND (1996), The Short Rate Diffusion Revisited: An Investigation Guided by the Efficient Method of Moments, *Workingpaper*.

ANDERSEN, T.G. AND J. LUND (1997), Stochastic Volatility and Mean Drift in the Short Rate Diffusion: Sources of Steepness, Level and Curvature in the Yield Curve, *Journal of Econometrics*.

BABBS, S.H. AND K.B. NOWMAN (1998), Kalman Filtering of Generalized Vasicek Term Structure Models, *Journal of Financial and Quantitative Analysis*, Forthcoming.

BALDUZZI, P., S.R. DAS, AND S. FORESI (1998), The Central Tendency: A Second Factor in Bond Yields, *Review of Economics and Statistics*, **80**, 62–72.

BALDUZZI, P., S.R. DAS, S. FORESI, AND R.K. SUNDARAM (1996), Stochastic Mean Models of the Term Structure of Interest Rates, *Workingpaper*.

BALTAGI, B.H. (1995), *Econometric Analysis of Panel Data*.

BAMS, D. (1997), An Empirical Comparison of Time Series and Cross Sectional Information in the Longstaff-Schwartz Term Structure Model, *Workingpaper Maastricht University*.

BAMS, D. AND P. SCHOTMAN (1997), A panel data analysis of affine term structure models, *LIFE Workingpaper 97-43*.

BEKAERT, G., R. HODRICK, AND D. MARSHALL (1997), On Biases in the Test of the Expectations Hypothesis of the Term Structure of Interest Rates, *Journal of Financial Economics*.

BERARDI, A. AND M. ESPOSITO (1997), A Base Model for Multifactor Specifications of the Term Structure, *Workingpaper, London Business School*.

BLISS, R.R. (1994), Testing Term Structure Estimation Methods, *Workingpaper Indiana University*, 1–39.

BLISS, R.R. AND P. RITCHKEN (1996), Empirical Tests of Two State-Variable Heath-Jarrow-Morton Models, *Journal of Money Credit and Banking*, **28**, 452–476.

BRENNAN, M.J. AND E.S. SCHWARTZ (1979), A Continuous Time Approach to the Pricing of Bonds, *Journal of Banking and Finance*, 133–155.

BRENNAN, M.J. AND E.S. SCHWARTZ (1992), An Equilibrium Model of Bond Pricing and a Test of Market Efficiency, *Journal of Financial and Quantitative Analysis*, 75–100.

BROWN, R.H. AND S.M. SCHAEFER (1994), The Term Structure of Real Interest Rates and the Cox, Ingersoll, Ross Model, *Journal of Financial Economics*, **35**, 3–42.

BROWN, S.J. AND P.H. DYBVIG (1986), The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates, *Journal of Financial Economics*, **41**, 617–630.

CAMPBELL, J.Y. (1986), A Defense of the Traditional Hypotheses about the Term Structure of Interest Rates, *Journal of Finance*, **41**, 183–193.

CAMPBELL, J.Y., A.W. LO, AND A.C. MACKINLAY (1997), *The Econometrics of Financial Markets*, Princeton University Press.

CHAMBERS, D.R., W.T. CARLETON, AND D.W. WALDMAN (1984), A New Approach to Estimation of the Term Structure of Interest Rates, *Journal of Financial and Quantitative Analysis*, **16**, 233–252.

CHAN, K.C., G.A. KAROLYI, F.A. LONGSTAFF, AND A.B. SANDERS (1992), An Empirical Comparison of Alternative Models for the Short-Term Interest Rate, *Journal of Finance*, **47**, 1209–1227.

CHEN, L. (1995), Stochastic Mean and Volatility : A Three-Factor Model of the Term Structure and its Application in Pricing of Interest Rate Derivatives, *Workingpaper FED Washington DC*.

CHEN, R.R. AND L. SCOTT (1992), Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure, *Review of Financial Studies*, **5**, 613–636.

CHEN, R.R. AND L. SCOTT (1993), Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates, *Journal of Fixed Income*, **3**, 14–31.

CORNELL, B. AND A.C. SHAPIRO (1989), The Mispricing of U.S. Treasury Bonds: A Case Study, *Review of Financial Studies*, **2**, 297–310.

COX, J., J. INGERSOLL, AND S. ROSS (1985), A Theory of the Term Structure of Interest Rates, *Econometrica*, **53**, 385–408.

COX, J., J. INGERSOLL, AND S. ROSS (1985), An Intertemporal General Equilibrium Model of Asset Prices, *Econometrica*, **53**, 363–384.

COX, J.C., J.E. INGERSOLL, AND S.A. ROSS (1981), A Reexamination of Traditional Hypotheses about the Term Structure of Interest Rates, *Journal of Finance*, **36**, 769–799.

CULBERTSON, J.M. (1957), The Term Structure of Interest Rates, *Quarterly Journal of Economics*, **71**, 485–517.

DAHLQUIST, M. AND L.E.O. SVENSSON (1994), Estimating the Term Structure of Interest Rates with Simple and Complex Functional Forms: Nelson & Siegel versus Longstaff & Schwartz, unpublished manuscript, Seminar paper 565, Institute of International Economic Studies, University of Stockholm.

DEJONG, F. (1997), Time-series and Cross-section Information in Affine Term Structure Models, *Workingpaper*.

DEMUNNIK, J. AND P.C. SCHOTMAN (1994), Cross Sectional versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market, *Journal of Banking and Finance*, **18**, 997–1025.

DERMODY, J.C. AND E.Z. PRISMAN (1988), Term Structure Multiplicity and Clientele in Markets with Transaction Costs and Taxes, *Journal of Finance*, **43**, 893–911.

DUAN, J.C. AND J.G. SIMONATO (1995), Estimating Exponential-Affine Term Structure Models by Kalman Filter, unpublished manuscript, Working paper, McGill University.

DUFFIE, D. (1996), *Dynamic Asset Pricing Theory*, Princton University Press, Princton New Jersey.

DUFFIE, D. AND R. KAN (1996), A Yield-Factor Model of Interest Rates, *Mathematical Finance*, **6**, 379–406.

DUFFIE, D. AND K. SINGLETON (1995), An Econometric Model of the Term Structure of Interest Rate Swap Yields, *Workingpaper Graduate School of Business, Stanford University*.

DYBVIG, P.H. (1989), Bond and Bond Option Pricing Based on the Current Term Structure, *Workingpaper*.

ELKAROUI, N., A. FRACHOT, AND H. GEMAN (1998), On the Behavior of Long Zero Coupon Rates in a No Arbitrage Framework, *Review of Derivatives Research*, **1**, 351–369.

ELKAROUI, N. AND V. LACOSTE (1992), Multifactor Models of the Term Structure of Interest Rates, *AFFI Conference Proceedings, Paris*.

ELKAROUI, N. AND J.C. ROCHET (1989), A Pricing Formula for Options on Coupon Bonds, *Workingpaper October, University of Paris VI*.

ESTRELLA, A. AND G.A. HARDOUVELIS (1991), The Term Structure as a Predictor of Real Economic Activity, *Journal of Finance*, **46**, 555–576.

EVANS, M.D.D. AND K.K. LEWIS (1994), Do Stationary Risk Premia Explain It All?, *Journal of Monetary Economics*, **33**, 285–318.

FAMA, E.F. (1984), Term Premiums and Default Premiums in the Money Markets, *Journal of Financial Economics*, **17**, 175–196.

FAMA, E.F. (1984), The Information in the Term Structure, *Journal of Financial Economics*, **13**, 529–546.

FAMA, E.F. AND R.R. BLISS (1987), The Information in Long Maturity Forward Rates, *The American Economic Review*, 680–692.

FRACHOT, A., D. JANCI, AND V. LACOSTE (1992), Factor Analysis of the Term Structure: A Probablistic Approach, *Workingpaper Banque de France*.

FRACHOT, A. AND J.P. LESNE (1993), Econometrics of Linear Factor Models of Interest Rates, *Workingpaper Banque de France*.

FRACHOT, A., J.P. LESNE, AND E. RENAULT (1995), Indirect Inference Estimation of Factor Models of the Yield Curve, *Workingpaper Banque de France*.

FRÜHWIRTH-SCHNATTER, S. AND A.L.J. GEYER (1997), Bayesian Estimation of Econometric Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure of Interest Rates

via MCMC Methods, unpublished manuscript, working paper, Vienna University of Economics and Business Administration.

GEYER, A.L.J. AND S. PICHLER (1997), A State-Space Approach to Estimate and Test Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure, unpublished manuscript, working paper, University of Economics, Vienna.

GIBBONS, M.R. AND K. RAMASWAMY (1993), A Test of the Cox, Ingersoll and Ross Model of the Term Structure, *The Review of Financial Studies*, 619–658.

GONG, F. AND E.M. REMOLONA (1996), A Three Factor Econometric Model of the US Term Structure, unpublished manuscript, Working Paper, Federal Reserve Bank New York.

GONG, F.F. AND E.M. REMOLONA (1997), Inflation Risk in the U.S. Yield Curve: the Usefulness of Indexed Bonds, *Workingpaper*.

HAMBURGER, M.J. AND E.N. PLATT (1975), The Expectations Hypothesis and the Efficiency of the Treasury Bill Market, *Review of Economics and Statistics*, **57**, 190–199.

HANSEN, L. (1982), Large Sample Properties of Generalized Method of Moment Estimators, *Econometrica*, 1029–1054.

HANSEN, L.P. AND R.J. HODRICK (1980), Forward Rates as Optimal Predictors of Future Spot Rates, *Journal of Political Economy*, **88**, 829–853.

HEATH, D., R. JARROW, AND A. MORTON (1992), Bond Pricing and the Term Structure of Interest Rates: A New Methodology, *Econometrica*, **60**, 77–105.

HESTON, S. (1991), Testing Continuous-Time Models of the Term Structure of Interest Rates, *Workingpaper School of Organization and Management, Yale University*.

HICKS, J.R. (1946), *Value and Capital*, London: Oxford University Press.

HSIAO, C. (1986), *Analysis of Panel Data*.

HULL, J. AND A. WHITE (1990), Pricing Interest Rate Derivative Securities, *The Review of Financial Studies*, **3**, 573–592.

HULL, J.C. (1993), *Options, Futures, and Other Derivative Securities*, Prentice Hall, Englewood Cliffs.

HULL, J.C. (1996), *Options, Futures, and other Derivatives*, Prentice Hall, 3^{rd} edition.

JAMSHIDIAN, F. (1989), An Exact Bond Pricing Formula, *Journal of Finance*, **44**, 205–209.

JAMSHIDIAN, F. (1991), Bond and Option Valuation in the Gaussian Interest Rate Model and Implementation, *Review of Financial Studies*, **9**, 131–170.

JAMSHIDIAN, F. (1992), A Simple Class of Square-Root Interest Rate Models, *Workingpaper Fuji Interenational London*.

JAMSHIDIAN, F. (1993), Bond, Futures and Option Evaluation in the Quadratic Interest Rate Model, *Workingpaper Fuji Interenational London*.

KAPPI, J. (1997), Testing Affine Yield Factor Models, *Journal of Financial Engineering*, **6**, ??

LANGETIEG, T. (1980), A Multivariate Model of the Term Structure of Interest Rates, *Journal of Finance*, **35**, 71–97.

LITTERMAN, R. AND J. SCHEINKMAN (1991), Common Factors affecting Bond Returns, *Journal of Fixed Incopme*, **1**, 54–61.

LONGSTAFF, F.A. AND E.S. SCHWARTZ (1992), A Two-Factor Interest Rate Model and Contingent Claim Valuation, *Journal of Fixed Income*, **4**, 16–23.

LONGSTAFF, F.A. AND E.S. SCHWARTZ (1992), Interest Rate Volatility and the Term Structure: a Two factor General Equilibrium Model, *Journal of Finance*, **47**, 1259–1282.

LONGSTAFF, F.A. AND E.S. SCHWARTZ (1993), Implemetation of the Longstaff-Schwartz Interest Rate Model, *Journal of Fixed Income*, **1**, 7–14.

LONGSTAFF, F.A. AND E.S. SCHWARTZ (1993), Interest Rate Volatility and Bond Prices, *Financial Analysis Journal*, **3**, 70–74.

LUND, J. (1997), Econometric Analysis of Continuous-Time Arbitrage-Free Models of the Term Structure of Interest Rates, unpublished manuscript, Working Paper, Aarhus Business School.

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

MCCULLOCH, J.H. (1975), The Tax-adjusted Yield Curve, *Journal of Finance*, **30(3)**, 811–830.

MELINO, A. (1988), The Term Structure of Interest Rates: Evidence and Theory, *Journal of Economic Surveys*, **2**, 335–366.

MISHKIN, F.S. (1988), The Information in the Term Structure: Some Further Results, *Journal of Applied Econometrics*, **3**, 307–314.

MODIGLIANI, F. AND R. SUTCH (1966), Innovations in Interest Rate Policy, *American Economic Review*, **56**, 178–197.

NELSON, C.R. AND A.F. SIEGEL (1987), Parsimonious Modeling of Yield Curves, *Journal of Business*, **60(4)**, 473–489.

PEARSON, N.D. AND T.S. SUN (1994), Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, Ross Model, *Journal of Finance*, **49**, 1279–1304.

PENNACHI, G. (1991), Identifying the Dynamics of Real Interest Rates and Inflation: Evidence Using Survey Data, *Review of Fiancial Studies*, **4**, 53–86.

RENAULT, E. (1996), Econometric Models of Option Pricing Errors, *Workingpaper GREMAQ*.

RONN, E.I. (1987), A New Linear Programming Approach to Bond Portfolio Management, *Journal of Financial and Quantitative Analysis*, **22**, 439–466.

SCHAEFER, S.M. AND E.D. SCHWARTZ (1984), A Two-Factor Model of the Term Structure: an Approximate Analytical Solution, *Journal of Financial and Quantitiative Analysis*, **19**, 413–424.

SCHOTMAN, P.C. (1996), A Bayesian Approach to the Empirical Valuation of Bond Options, *Journal of Econometrics*, **75**, 183–216.

SERCU, P. AND X. WU (1997), The Information Content in Bond Model Residuals: An Empirical Study on the Belgian Bond Market, *Journal of Banking and Finance*, **21**, 685–720.

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

STAMBAUGH, R. (1988), The Information in Forward Rates: Implications for Models of the Term Structure, *Journal of Financial Economics*, **21**, 41–70.

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

VASICEK, O.A. AND H.G. FONG (1982), Term Structure Modeling Using Exponential Splines, *The Journal of Finance*, **37(2)**, 339–348.

WHITE, H. (1980), A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica*, **48**, 817–838.

Curriculum Vitae

Dennis Bams was born on June 18^{th}, 1970 in Maastricht, the Netherlands. He studied Econometrics and Operations Research at the Erasmus University in Rotterdam and Actuarial Science at the University of Amsterdam. In 1994 he finished his Master's thesis, *"Asset Liability Management for Pensionfunds: a Multistage Stochastic Programming Approach"*. For two years he worked as a quantitative researcher for Pacific Investment Holding in Amsterdam and Ortec Consultants in Gouda. In 1994, Dennis Bams, moved to Maastricht University to work for the Finance Department where he wrote this Ph.D. dissertation on the term structure of interest rates. Part of this research was done while he was a visiting sholar at GREQAM, Ecole des Hautes Etudes en Sciences Sociales in Marseille, France, in the summer of 1996. The first half of 1997 he spent as a participant of the Ph.D. program at the Graduate School of Business at the University of Chicago. Since October 1^{st}, 1998, he has been working at the ALM Insurance Department of ING Bank in Amsterdam.

Dennis Bams is op 18 juni 1970 in Maastricht geboren. Hij studeerde Econometrie en Operations Research aan de Erasmus Universiteit te Rotterdam, en Actuariële Wetenschappen aan de Universiteit van Amsterdam. In 1994 schreef hij zijn afstudeerscriptie, *"Asset Liability Management for Pensionfunds: a Multistage Stochastic Programming Approach"*. Gedurende twee jaren werkte hij als een kwantitatieve onderzoeker bij Pacific Investment Holding te Amsterdam en voor Ortec Consultants te Gouda. In 1994, begon Dennis Bams te werken voor sectie financiering aan de Universiteit Maastricht. Daar schreef hij zijn proefschrift over de rente termijnstructuur. Een gedeelte van het onderzoek is uitgevoerd aan GREQAM, Ecole des Hautes Etudes en Sciences Sociales in Marseille, Frankrijk, in de zomer van 1996. De eerste helft van 1997 hij heeft doorgebracht aan de Graduate School of Business van de University of Chicago. Vanaf 1 october 1998 is hij werkzaam bij de afdeling ALM Verzekeringen van ING Bank te Amsterdam.

The term structure of interest rates plays a key role in many economic developments. It directly determines the value of fixed income securities, like bonds, and it heavily influences the attractiveness of equities. Not only is the term structure of interest rates important for large institutional investors, it also has important consequences for individuals, who have to decide upon the type of mortgage contract, or whether it is interesting to buy a life insurance policy. The importance of the term structure of interest rates leads to the empirical analysis of various interest rate models in this dissertation. We carefully set up an econometric framework to investigate the properties of the affine class of term structure models. We pursue a panel data analysis, and compare the results for both endogenous and exogenous models. The results are compared with time series approaches and also the implications for other interest rate derivative securities have been investigated.

Dennis Bams was born on June 18th, 1970 in Maastricht, the Netherlands. He studied Econometrics and Operations Research at the Erasmus University in Rotterdam and Actuarial Science at the University of Amsterdam. In 1994 he finished his Master's thesis, *Asset Liability Management for Pensionfunds: a Multistage Stochastic Programming Approach*. For two years he worked as a quantitative researcher for Pacific Investment Holding in Amsterdam and Ortec Consultants in Gouda. In 1994, Dennis Bams moved to Maastricht University to work for the Finance Department where he wrote this Ph.D. dissertation on the term structure of interest rates. Part of this research was done while he was a visiting scholar at GREQAM, Ecole des Hautes Etudes en Sciences Sociales in Marseille, France, in the summer of 1996. The first half of 1997 he spent as a participant of the Ph.D. program at the Graduate School of Business at the University of Chicago. Since October 1st, 1998, he has been working at the ALM Insurance Department of ING Bank in Amsterdam.

ING GROEP

ISBN 90 9012476 4