ELSEVIER

Journal of Empirical Finance 8 (2001) 297–323

Journal of EMPIRICAL FINANCE

www.elsevier.com/locate/econbase

The joint estimation of term structures and credit spreads

Patrick Houweling [a,*], Jaap Hoek [b], Frank Kleibergen [c]

[a] *Central Market Risk, Rabobank International and Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, H16-20, 3000 DR Rotterdam, The Netherlands*

[b] *RObeco Center of Knowledge, Robeco Group, The Netherlands*

[c] *Department of Quantitative Economics, University of Amsterdam, The Netherlands*

Accepted 21 March 2001

Abstract

We present a new framework for the joint estimation of the default-free government term structure and corporate credit spread curves. By using a data set of liquid, German mark denominated bonds, we show that this yields more realistic spreads than traditionally obtained spread curves that result from subtracting independently estimated government and corporate term structures. The obtained spread curves are smooth functions of time to maturity, as opposed to the twisting curves one gets from the traditional method, and are less sensitive to model specifications. To determine the 'optimal' model specification, we use a newly developed test statistic that compares spread curves from competing models.

JEL classification: G12; G13; C13

Keywords: Term structure estimation; Credit spreads; Corporate bonds; Splines

1. Introduction

In the 1990s we witnessed a wave of new models that take credit risk in financial instruments into account. Catalysing factors that spurred research in this

* Corresponding author. Tel.: +31-10-408-8938; fax: +31-10-408-9031.

E-mail addresses: houweling@few.eur.nl (P. Houweling), j.hoek@robeco.nl (J. Hoek), kleiberg@fee.uva.nl (F. Kleibergen).

0927-5398/01/$ - see front matter
PII: S0927-5398(01)00026-3

area were the growth of the corporate bond market, the rapid development of the market for credit derivatives and the growing awareness among investors of credit risk in derivative products. Many of the new models require an accurate description of the term structures of interest rates of different credit risk classes as input data. Measuring a term structure for a particular credit rating class amounts to estimating its credit spread curve relative to the government curve, which proxies the default-free curve. Traditionally, spread curves are calculated by subtracting independently estimated government and corporate term structures. In this paper we present a new framework that jointly estimates the government curve and credit spread curves. Unlike the twisting curves one gets from the traditional method, the estimated spread curves are now smooth functions of time to maturity, and are less sensitive to model settings. Moreover, we develop a novel test statistic that allows us to determine the 'optimal' settings of the new model.

An important application in which accurately estimated term structures of interest rates are essential inputs is the pricing of defaultable bonds and credit derivatives. The leading frameworks are the Jarrow et al. (1997) Markov chain model, which extended the work of Litterman and Iben (1991) and Jarrow and Turnbull (1995) to multiple credit ratings, and the Duffie and Singleton (1999) framework, which can be cast into a defaultable Heath et al. (1992) model. Similar to the default-free interest rate models developed in the early 1990s—most notably the extended Vasicek (1977) models, such as Hull and White (1990), the lognormal short rate models, like Black et al. (1990), and the models in the Heath et al. (1992) framework—these credit risk models provide an exact fit to today's default-free and defaultable term structures of interest rates. Any error in the input of such models will be amplified in the prices of interest rate and credit derivatives that are subsequently priced with them.

Interest rates and spread curves are also required for risk management purposes, for example, in applying the historic simulation method to calculate the Value at Risk for a corporate bond portfolio (see, e.g. Saunders, 1999, Chapter 11). Future scenarios are generated by adding historical day-to-day movements in interest rates and spread curves to today's curves. Since in each scenario the bond portfolio is revalued to obtain the empirical distribution of the future portfolio value, inaccurate curves may lead to an unnecessarily large Value at Risk and a too large amount of regulatory capital. Other applications of default-free and defaultable interest rates include the pricing of new bond issues and assessing counterpart risk in derivative products (see, e.g. Hull and White, 1995; Duffee, 1996; Caouette et al., 1998, Chapter 21).

An obstacle in the above mentioned applications is that the term structures are not directly observable in the market and have to be estimated from market prices of traded instruments using statistical techniques. Until now, the literature has primarily focused on the estimation of the default-free term structure from a sample of government bonds. The standard approach originates from McCulloch (1971, 1975), who modelled the discount curve as a linear combination of polynomial basis functions. Other approaches include the use of Bernstein polynomials (Schaefer, 1981), exponential splines (Vasicek and Fong, 1982), *B*-splines (Shea, 1985; Steeley, 1991), exponential forms (Nelson and Siegel, 1987) or a bootstrapping procedure as employed on electronic information systems Bloomberg and Reuters; Anderson et al. (1996, Chapter 2) provided an extensive overview of these and other term structure estimation methods. After choosing one of these methods, we could independently estimate a separate model for each credit class. We illustrate that these calculations are likely to result in twisting spread curves that alternately have positively and negatively sloped segments. Moreover, the level and shape of the spread are shown to be sensitive to model misspecification.

Instead, we jointly estimate the default-free and defaultable interest rate curves. Our joint estimation is based on the decomposition of a defaultable term structure into a default-free curve and a credit spread curve. The default-free curve is estimated from government bonds, so that our model for a corporate term structure focuses on the credit spread only. Both the government curve and the corporate spread curve are modelled as *B*-spline functions and all parameters are jointly estimated from a combined data set of bonds. We apply the model to a data set of liquid, German mark denominated bonds, whose credit ratings range from Standard and Poor's ratings AAA to B. We obtain smooth and reliably estimated spread curves that are relatively robust to model misspecification. Moreover, we demonstrate that these results can be attributed to both the *joint* and the *parsimonious* modelling. Independently estimating the government curve and a parsimoniously specified corporate curve model does not yield the same results, nor does jointly estimating the government curve and a richly specified corporate spread curve.

The remainder of the paper is structured as follows. Section 2 presents the new framework for the joint estimation of the government term structure and corporate credit spread curves. The specification of the model is described in Section 3, whereas Section 4 goes over several methods to choose between competing models, including a novel statistic that is developed to compare spread curves obtained from alternative model specifications. Section 5 applies the new model to our data set and confronts the performance of jointly estimated term structures with independently estimated term structures. Section 6 concludes the paper.

2. Multi-curve model

Ideally, we would like to use a different spread curve for each firm, reflecting the uniqueness of a firm's characteristics that determine its credit risk. Due to data constraints, however, which are discussed in Section 5, we have to resort to grouping firms that have similar credit worthiness and face similar operating environments. A disadvantage of grouping bonds is that a particular type of heterogeneity may occur (Helwege and Turner, 1999).[1] Suppose we have created C categories of bonds, where category 1 corresponds to government bonds and the other categories are formed by using, e.g. credit rating and industry as criteria. The purpose is to estimate a spread curve for each category. Instead of independently estimating term structures, we propose a joint estimation approach. Since a corporate term structure consists of a default-free curve and a credit spread curve, it seems natural to only model the spread and take the default-free part from the government curve. Several representations of the term structure exist, e.g. as discount factors or spot interest rates, but it is common practice to model the discount curve. We use the following framework for jointly estimating the discount curves:

$$D_1(t) = d(t)$$
$$D_c(t) = d(t) + s_c(t), \quad c = 2,3,\ldots,C, \tag{1}$$

where $D_c(\cdot)$ is the discount curve of category c, $d(\cdot)$ is the model for the government discount curve and $s_c(\cdot)$ is the model for the discount spread curve of category c with respect to the government curve. We impose C constraints $D_c(0) = 1$, because a payment due today does not need to be discounted. All parameters in the models for the government curve and the discount spread curves are jointly estimated from a combined data set of government and corporate bonds. We refer to this model as the *multi-curve model* as opposed to a *single-curve model* that independently estimates a single corporate term structure.

To model $d(\cdot)$ and $s_c(\cdot)$, we use *spline functions*, as introduced to the term structure estimation literature by McCulloch (1971). Some commonly used spline models are the exponential splines, by Vasicek and Fong (1982), and the B-splines, as discussed by Shea (1985) and Steeley (1991). Splines are tailored to approximate a scatter of data points by a continuous and preferably smooth function. Their main advantage is their flexibility: there is no need to a priori impose a particular curvature, because the shape of the curve is determined by the data. Bliss (1997) compared several non-parametric term structure estimation models and found that spline models perform at least as good as competing models, and outperform the other considered models if the data contains longer maturity bonds (over 5 years).

Splines are basically piecewise polynomials. The approximation interval[2] $[a, b]$ is divided into n subintervals $[\tau_0, \tau_1], [\tau_1, \tau_2], \ldots, [\tau_{n-1}, \tau_n]$, where the knots τ_i are chosen such that $a = \tau_0 < \tau_1 < \ldots < \tau_n = b$. The data points in each subinter-

[1] Within a data set of bonds of the same rating, the longest maturity bonds usually have been issued by the relatively most credit worthy firms. Therefore, credit spreads may decrease for the longest maturities in such a data set. We are grateful to an anonymous referee for bringing this to our attention.

[2] With term structure estimation the approximation interval runs from 0 to the longest bond maturity in the sample.

val are modelled as a kth degree polynomial. The n polynomials are constrained by the condition that the spline has to be $k-1$ times continuously differentiable. This is a restriction at the knots $\tau_1, \tau_2, \dots, \tau_{n-1}$ only, and imposes k constraints on the coefficients of two adjacent polynomials. In sum, we have $n(k+1)$ coefficients minus $(n-1)k$ constraints, leaving only $n+k$ degrees of freedom. A more parsimonious way of representing splines is by means of basis functions; see, e.g. Powell (1981, p. 228). Any kth degree spline function $S(\cdot)$ with knots $\boldsymbol{\tau}$ can be expressed as a linear combination of $n+k$ basis functions $\mathbf{f}(\cdot) = \{f_1(\cdot), f_2(\cdot), \dots, f_{n+k}(\cdot)\}$:

$$S(t) = \sum_{s=1}^{n+k} \alpha_s f_s(t) = \mathbf{f}(t)' \boldsymbol{\alpha}.$$

Once the basis is chosen and the degree k and the knots $\boldsymbol{\tau}$ are set, the basis functions are fully specified. The spline weights $\boldsymbol{\alpha}$, however, are unknown and have to be estimated from the data. Powell (1981) recommended the use of a basis of B-spline functions, because of their efficiency and numerical stability. Steeley (1991) applied B-splines to term structure estimation. See Appendix A for a concise description of constructing a basis of B-splines, or consult Powell (1981) for more details.

We use B-splines to model the government discount curve and corporate discount spread curves in Eq. (1). We set $d(t) = \mathbf{g}_1(t)'\beta_1$ and $s_c(t) = \mathbf{g}_c(t)'\beta_c$, where $\mathbf{g}_i(\cdot)$ contains $n_i + k_i$ B-spline basis functions that span a spline of k_ith degree with knots $\boldsymbol{\tau}_i$; Section 3 discusses the specification of the degrees and knots. Using B-spline basis functions, the multi-curve model (Eq. (1)) is rewritten as:

$$D_1(t) = \mathbf{g}_1(t)' \beta_1 \tag{2a}$$

$$D_c(t) = \mathbf{g}_1(t)' \beta_1 - \mathbf{g}_c(t)' \beta_c, \quad c = 2,3,\dots,C. \tag{2b}$$

To estimate the unknown spline weights $\beta_1, \beta_2, \dots, \beta_C$, we construct a data set, consisting of B_c bonds of category c, and use the *discounted cash flow* (DCF) principle to link the bond prices to a discount curve. According to the DCF principle, the price that an investor is willing to pay for the bth bond of category c equals the sum of the present values of the cash flows:

$$P_{\mathrm{DCF},cb} = \sum_{i=1}^{N_{cb}} \mathrm{CF}_{cbi} D_c(t_{cbi}), \tag{3}$$

where $P_{\mathrm{DCF},cb}$ is the DCF bond price, N_{cb} is the number of remaining cash flows and CF_{cbi} is the ith cash flow that is paid at time t_{cbi}. By using the DCF equation as the theoretical bond price model, we have to confine our data set to fixed-income bonds with known redemptions and exclude any bonds with optional elements—such as callable and puttable bonds—and bonds with floating or index-linked coupons. The DCF method is valid if we assume a perfectly competitive capital market, i.e. if all relevant information is widely and freely available and no barriers, frictions and taxes exist. Brealey and Meyers (1991, p. 20) stated that "even though these conditions are not fully satisfied, there is considerable evidence that security prices behave almost as if they were."

For category 1, i.e. for government bonds, we substitute Eq. (2a) into Eq. (3), yielding,

$$P_{\mathrm{DCF},1b} = \sum_{i=1}^{N_{1b}} \mathrm{CF}_{1bi} \left(\sum_{s=1}^{n_1+k_1} \beta_{1s} \mathbf{g}_{1s}(t_{1bi}) \right) = \sum_{s=1}^{n_1+k_1} \beta_{1s} \left(\sum_{i=1}^{N_{1b}} \mathrm{CF}_{1bi} \mathbf{g}_{1s}(t_{1bi}) \right)$$
$$\equiv \sum_{s=1}^{n_1+k_1} \mathbf{x}_{1bs} \beta_{1s} = \mathbf{x}'_{1b} \beta_1, \tag{4a}$$

whereas for categories 2, 3, ..., C, substitution of Eq. (2b) into Eq. (3) results in,

$$P_{\mathrm{DCF},cb} = \sum_{s=1}^{n_1+k_1} \beta_{1s} \left(\sum_{i=1}^{N_{cb}} \mathrm{CF}_{cbi} \mathbf{g}_{1s}(t_{cbi}) \right) + \sum_{s=1}^{n_c+k_c} \beta_{cs} \left(\sum_{i=1}^{N_{cb}} \mathrm{CF}_{cbi} \mathbf{g}_{cs}(t_{cbi}) \right)$$
$$\equiv \sum_{s=1}^{n_1+k_1} \mathbf{x}_{cbs} \beta_{1s} + \sum_{s=1}^{n_c+k_c} \mathbf{y}_{cbs} \beta_{cs} = \mathbf{x}'_{cb} \beta_1 + \mathbf{y}'_{cb} \beta_c. \tag{4b}$$

Note that these equations for the theoretical bond price are linear in the unknown parameters, because all terms in $\mathbf{x}$ and $\mathbf{y}$ are either known from the characteristics of the bond or the specification of the spline models. Also, the constraints $D_c(0) = 1$ on the discount functions are linear restrictions on the spline weights; see Eqs. (2a) and (2b).

In order to estimate the spline weights, we substitute the theoretical prices $P_{\mathrm{DCF},cb}$ by observed market prices P_{cb} and add an error term ε_{cb} to the equations. The error term is necessary, because due to market imperfections the DCF method is not able to perfectly explain bond prices.[3] Using matrix notation, we obtain the following linear regression model:

$$\begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \\ \vdots \\ \mathbf{P}_C \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & 0 & 0 & \cdots & 0 \\ \mathbf{X}_2 & \mathbf{Y}_2 & 0 & \cdots & 0 \\ \mathbf{X}_3 & 0 & \mathbf{Y}_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{X}_C & 0 & 0 & \cdots & \mathbf{Y}_C \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \boldsymbol{\beta}_3 \\ \vdots \\ \boldsymbol{\beta}_C \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \boldsymbol{\varepsilon}_3 \\ \vdots \\ \boldsymbol{\varepsilon}_C \end{bmatrix}, \quad \varepsilon_C \sim \text{i.i.d.}\left(0, \sigma_c^2\right), \tag{5}$$

where $\mathbf{X}_c$ is a $B_c \times (n_1 + k_1)$ matrix with rows $\{\mathbf{x}'_{c1}, \mathbf{x}'_{c2}, \ldots, \mathbf{x}'_{cB_c}\}$ and $\mathbf{Y}_c$ is a $(B_c \times (n_c + k_c))$ matrix with rows $\{\mathbf{y}'_{c1}, \mathbf{y}'_{c2}, \ldots, \mathbf{y}'_{cB_c}\}$. We allow the disturbances

[3] It is possible to obtain an arbitrary high goodness of fit by increasing the number of parameters. However, the resulting term structures are likely to have twisting shapes and wide confidence intervals.

to have different variances for each category, because prices of lower rated bonds are generally more noisy due to lower liquidity and a higher uncertainty about their perceived credit worthiness. Also, the residuals of independently estimated single-curve models can be shown to have significantly different variances using a heteroscedasticity test. Estimates $\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \ldots, \hat{\boldsymbol{\beta}}_C$ of the spline weights are readily obtained by applying Restricted Feasible Generalised Least Squares estimation to Eq. (5); see, e.g. Greene (2000, p. 473). Once we have estimated the spline weights, we can evaluate $\hat{D}_c(t)$ for any maturity t. It is important to emphasise, however, that discount factors for maturities beyond the maximum maturity bond become unreliable.

3. Model settings

Before we are able to estimate the regression model in Eq. (5), we have to specify the exact form of the basis functions. The functional form of the basis functions follows by choosing the degree of the splines and the number and location of the knots. These choices reflect the familiar trade-off between flexibility and smoothness. The degree of the splines should not be chosen too high, to preclude the problems of higher order polynomials.[4] If the order is too low, however, the estimated curve will not fit the data very well, and thus will not reflect the interest rates that are prevalent in the market. Similarly, if the number of knots is chosen too low, the model will not be able to fit term structures with difficult shapes. On the other hand, if it is too high, the estimated curve is sensitive to outliers.

For the spline model for the government discount curve, we can use results from the term structure estimation literature. Almost all studies that employ spline functions to model the discount curve use third degree splines. Only McCulloch (1971) used quadratic splines in his pioneering study, resulting in 'knuckles' in the forward curve, which made him switch to cubic splines in his follow-up paper (McCulloch, 1975). Beim (1992) conducted a simulation study and concluded that cubic splines are preferable. Poirier (1976, p. 49) demonstrated that fitting a cubic spline minimises the integral of the square of the second derivative, which is an approximation of a function's smoothness; see also Adams and van Deventer (1994). Consequently, cubic splines are a convenient compromise between high goodness of fit and smooth curves.

[4] Using higher order polynomials often results in spurious curvature between the data points, see, e.g. Shea (1984, p. 255). This is especially true if such polynomials are fitted to data that are not uniformly distributed over the approximation interval, as is the case with term structure estimation, see Fig. 1 in Section 5.

With regard to the specification of the knots, there is less agreement. McCulloch (1971) proposed to set the number of knots equal to the integer nearest to the square root of the number of bonds in the sample. The knots are then located such that an approximately equal number of bonds is placed in each segment. Litzenberger and Rolfo (1984) stated that the McCulloch scheme is likely to result in a poor fit for longer maturities due to the larger number of shorter maturity bonds. As an alternative, they suggested to exogenously place the knots at 1, 5 and 10 years, roughly corresponding to an economic segmentation into short, medium and long maturities. Langetieg and Smoot (1989) tested the McCulloch knot placement scheme against the economic scheme and found that the latter typically performed better. Steeley (1991) experimented with the specification of the knots and recommended placing knots at 5 and 10 years as a starting point for future research. The simulation study by Beim (1992) revealed that cubic splines with two knots minimised the standard error of fit between the estimated and the simulated 'true' discount curves.

We also use spline functions to model discount spread curves, but there is no prior evidence available on the specification of the degree and the knots. Given the disagreement in the literature on the specification of the default-free discount curve, our task of specifying the splines of the spread curve is not an easy one. Our choices are guided by the observation that a spread curve generally has a less complicated shape than a term structure. Therefore, we reduce the flexibility of the spline model for the spread curve by reducing the degrees of freedom. Compared to the spline for the discount curve of category c in a single-curve model, we specify the discount spread curve $s_c(\cdot)$ in the multi-curve model as a lower degree spline with a smaller number of knots. That is, we use a quadratic spline function and the knots are chosen to be a subset of the knots of the single-curve spline model. This still leaves us with several competing degree–knot combinations. To choose the 'optimal' combination, we use a newly developed test statistic that allows us to compare spot spread curves that are obtained from competing multi-curve models. We describe this *curve similarity test* in the next section and apply it in Section 5.

4. Model comparison

A problem in comparing different single-and multi-curve models among and against one another is that there is no general estimable model that encompasses all other models. Therefore, we cannot use standard econometric testing procedures. Moreover, most econometric tests only focus on goodness of fit, i.e. the ability of a model to fit the data. In term structure estimation, however, practitioners are additionally interested in other features of the models, such as smoothness and statistical reliability. For these reasons, we compare single-and multi-curve models in three ways.

(a) Usage of statistics that reflect the goodness of fit, smoothness and reliability, such that models can be compared by confronting these statistics, though without being able to determine the statistical significance of possible differences.

(b) Usage of a newly developed test statistic that allows two curves from two different multi-curve models to be compared to one another. We focus here on spot spread curves, because of their importance as inputs for pricing and risk management models, but the statistic may also be employed to compare other curves that can be calculated from the multi-curve models.

(c) Visual inspection of the estimated term structures, most notably the spot curves and the spot spread curves. Desirable features are smoothness and monotonicity.

Issues a and b will be discussed in more detail in the remainder of this section.

4.1. Statistics

Since interest rates are the main determinants of bond prices, any term structure model should be able to explain market prices fairly accurately. Therefore, *goodness of fit* is a useful criterion to compare models. We measure the fit as the Root Mean Squared Error of the residuals:

$$\mathrm{RMSE}_c = \left(\frac{1}{B_c} \sum_{b=1}^{B_c} e_{cb}^2 \right)^{1/2},$$

where RMSE_c denotes the Root of Mean Squared Error for category c, B_c the number of category c bonds and e_{cb} the residual of the bth category c-bond, which is calculated as the market price of the bond minus its theoretical DCF price (4a) or (4b).

Although a low value of the RMSE statistic is desirable, we run the risk of ending up with a twisting curve. Therefore, we also measure the smoothness of estimated term structures. Following Poirier (1976, p. 49) and Powell (1981), the smoothness of a function ϕ over an interval $[t_1, t_2]$ is computed as the integral of the square of its second derivative:

$$s(\phi, t_1, t_2) = \int_{t_1}^{t_2} \phi''(t)^2 \, \mathrm{d}t.$$

We evaluate this statistic for both spot curves and spot spread curves.

Finally, we want to judge to what degree deviations between theoretical and market prices are transformed into uncertainty about estimated interest rates and spreads. The *reliability* of a point (maturity) on an estimated curve is indicated by its standard error. The reliability can be evaluated in a number of maturities to compare different segments of the curve. Appendix B derives the standard errors for a number of curves: discount curve, discount spread curve, spot curve and spot spread curve.

4.2. Curve similarity test

The Curve Similarity Test described below helps in choosing between two curves that are estimated with two different multi-curve models. The test especially guides in striking a balance between goodness of fit and smoothness (see Section 3). Given our focus on credit risk models, we describe the construction of the test for spot spread curves, but the test is suitable for any other curve for which standard errors and covariances can be computed.

Suppose we estimate a richly specified multi-curve model and compose a vector $\hat{\mathbf{s}}^1_{r,c}(\mathbf{t})$ of spot spread rates of category c evaluated in a q-vector of maturities $\mathbf{t} = (t_1, \dots, t_q)$. We would like to know to what extent we can reduce this model to a more parsimonious model, i.e. a smoother curve, without loosing too much on the goodness of fit criterion. Consider therefore the vector $\hat{\mathbf{s}}^0_{r,c}(\mathbf{t})$ that contains spot spread rates—evaluated in the same vector of maturities—that result from a more parsimonious multi-curve model. This alternative model contains less parameters due to a lower degree and/or less spline knots. The Curve Similarity Test (CST) aims to test whether $\hat{\mathbf{s}}^0_{r,c}(\mathbf{t})$ lies in the realm of $\hat{\mathbf{s}}^1_{r,c}(\mathbf{t})$.

To compute the CST statistic we weigh differences between the spot spread vectors with the covariance matrix $\Sigma^1_c(\mathbf{t})$ of $\hat{\mathbf{s}}^1_{r,c}(\mathbf{t})$:

$$\text{CST} = \left(\hat{\mathbf{s}}^1_{r,c}(\mathbf{t}) - \hat{\mathbf{s}}^0_{r,c}(\mathbf{t})\right)' \left(\Sigma^1_c(\mathbf{t})\right)^{-1} \left(\hat{\mathbf{s}}^1_{r,c}(\mathbf{t}) - \hat{\mathbf{s}}^0_{r,c}(\mathbf{t})\right).$$

The covariance matrix, which is constructed in Appendix B, measures the uncertainty in the spread estimates, and by using it as weight matrix, we put more emphasis on the reliable maturities of the spread curve, and vice versa. We compare the CST value to critical values from a χ^2 distribution with q degrees of freedom to determine whether $\hat{\mathbf{s}}^0_{r,c}(\mathbf{t})$ is approximately equal to $\hat{\mathbf{s}}^1_{r,c}(\mathbf{t})$ at the selected maturities. The testing procedure can only be applied to spot spread vectors from multi-curve models. Spread curves from single-curve models are obtained by subtracting independently estimated term structures, so that we are unable to construct the covariance matrix of a spread vector. As $\hat{\mathbf{s}}^1_{r,c}(\mathbf{t})$ curve we choose the multi-curve model with the same degree–knot settings as the single-curve model, because its spread curve resembles the spread curves obtained from single-curve models the most (see Section 5). As $\hat{\mathbf{s}}^0_{r,c}(\mathbf{t})$ curves we consider several more parsimonious multi-curve models, i.e. with a lower degree and/or less knots. These different parsimonious models are all compared to the most richly specified model. The results that stem from such a model comparison should be interpreted with care as the testing procedure is conceptually different from standard econometric testing procedures. For example, the test statistic may prefer a model with low order splines that has appropriately selected knots to a high-order model with badly located knots. Therefore, the test may reject a model that has a larger number of parameters than a competing model that is not rejected; this outcome is not possible with traditional econometric tests that compare nested models.

To make the test operational, we have to specify the maturity vector $\mathbf{t}$. Since the covariance matrix of the spot spread vector $\hat{\mathbf{s}}^1_{r,c}(\mathbf{t})$ is derived from the covariance matrix of the estimators $(\hat{\boldsymbol{\beta}}^1_1,\ \hat{\boldsymbol{\beta}}^1_c)$ of the richly specified model, we cannot construct the covariance matrix for an arbitrarily chosen maturity vector $\mathbf{t}$. For example, if the number of maturities q exceeds the number of parameters in the underlying regression model, the covariance matrix of $\hat{\mathbf{s}}^1_{r,c}(\mathbf{t})$ becomes singular. Another issue is the location of the maturities. Because of the smoothness of the curve, spreads for two adjacent maturities cannot be very different from each other. Therefore, the grid points should not be chosen too close to each other to preclude a near-singular covariance matrix. A final consideration is the location of the maturities relative to the spline knots. Since each spline interval corresponds to an extra parameter, we cannot place too much grid points of the test in one spline interval. Again, doing so would lead to a near-singular matrix. In practice, the above mentioned conditions on the maturity vector $\mathbf{t}$ imply that we can only conduct a joint comparison of the spot spreads at a limited number of maturities, which lie reasonably far apart. To determine the robustness of the results from the testing procedure, we can vary the maturity vector $\mathbf{t}$ while satisfying the conditions.

5. Results

5.1. Data

To appraise the performance of the proposed multi-curve model and compare it to independently estimated single-curve models, we use a data set of German mark (DEM) denominated bonds. Their characteristics like maturity dates, coupon percentages and credit ratings, are obtained from Bloomberg, whereas bond quotes are retrieved daily at 4:00 pm from Reuters' TREASURY and EUROBOND pages. These Reuters pages are connected to broker pages, and each time a broker updates a quote for a bond, that quote is also refreshed on the Reuters page. Therefore, the TREASURY and EUROBOND pages provide a good representation of the market for German mark denominated bonds.

For illustration, we present the results for the trading days of June 1998, yielding a total of 1291 quoted bonds. To estimate the term structures, we construct a sample of fixed-coupon, bullet bonds, and to ensure their liquidity, we only use bonds that are quoted on at least 18 of the 20 trading days of June 1998. There are 624 bonds that satisfy these conditions. In estimating the term structure on a particular trading day, we also consistently exclude all bonds with a remaining maturity of less than 1 year. Unlike the US and UK Treasury Bills, short-term German government bonds and short-term discount bonds of other credit ratings typically have low liquidity. In our data set, such bonds showed constant prices or they were not quoted at all for several consecutive days.

Since we cannot use short-term bonds, we add four synthetic zero-coupon bonds to the sample, whose prices correspond to 1-, 3-, 6- and 12-month money-market rates, respectively (see also Bühler et al., 1999). We use Frankfurt Interbank Offered Rates (FIBORs) as money-market rates, but since most commercial banks have an AA rating, these rates are not straightly applicable to other rating classes. Therefore, we correct the FIBORs by adding or subtracting a category-and maturity-dependent spread. The price P_{ct} of a synthetic bond for category c of maturity t is thus computed as:

$$P_{ct} = \frac{1}{(1 + \mathrm{FIBOR}_t + \mathrm{correction}_{ct})\,t}.$$

Note that we do not constrain the curves to pass exactly through the corrected FIBORs; the synthetic bonds are just additional data points to support the curve in a sparse data segment. Due to the subjectivity of corrections, however, care must be taken in using rates from the short end of the estimated term structures.

The number of suitable bonds of a single corporate debtor is too small to reliably estimate a separate term structure for each debtor. Our data set comprises 168 unique issuers; only 14 of them have 10 or more suitable bonds outstanding, and only 2 out of these 14 issued more than 20 suitable bonds. Therefore, we resort to grouping firms by rating and industry. First we show a division of the included bonds by rating; see Table 1. We use ratings published by Bloomberg, which compounds the major ratings of rating agencies Moody's and Standard and Poor's. We consider AAA-rated government bonds as a separate category, indicated by 'rating' symbol GOVT. From the table, it is clear that the number of bonds per rating decreases with credit quality. Therefore, the reliable estimation of term structures of lower rating categories may be hampered by an insufficient number of bonds if we were to use a single-curve method. The proposed multi-curve model offers a solution, because it only focuses on the spread curve and involves less parameters.

The distribution of the bond's maturity dates is shown in Fig. 1. For government bonds and AAA bonds, large gaps in the maturity distribution are observed beyond a maturity of 10 years. This is caused by the very infrequent issuance of

Table 1
Distribution of bonds in the data set by rating
All are all German Mark Eurobonds that are quoted at least once in June 1998. *Included* are fixed-income, bullet bonds that are quoted on at least 18 days of the 20 trading days of June 1998. Rating symbols: *GOVT*: AAA-rated government bonds; *NR*: not rated; *AAA, AA, A, BBB, BB, B*: Standard and Poor's major rating letters.

	GOVT	AAA	AA	A	BBB	BB	B	NR	Total
All	112	411	329	119	30	53	38	199	1291
Included	93	228	145	53	13	16	16	60	624

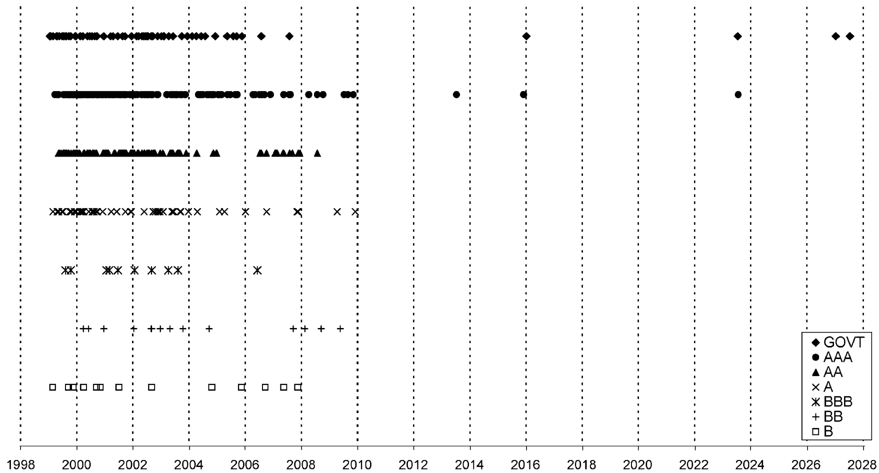

Fig. 1. Distribution of the maturity dates of the bonds in the estimation sample by rating category.

long-term bonds (e.g. 30-year bonds), as opposed to the regular issuance of 5- and 10-year bonds. Therefore, a term structure estimation procedure has few data points beyond 10 years and identification of that part of the term structure is more difficult.

For credit classes AAA and AA, there are enough bonds to allow a further classification by industry, but for the other ratings this is not feasible due to the limited number of bonds. The included bonds from ratings AAA and AA are assigned to one of four compounded Reuters industry classifications. According to Table 2, the majority of the bonds is issued by financial institutions, whereas supra-national organisations are also well represented.

5.2. *Single-curve results*

Before we discuss the results of the multi-curve (MC) model, we first present some estimation results of single-curve (SC) models. We only describe the results

Table 2

Distribution of 'included' bonds in the data set by rating and industry

Financials consists of Reuters classifications 'Banks' and 'Financials', *Government Agencies* is made up by 'Government National', 'Government Regional' and 'Government Agency', *Supra-National* consists of 'Worldbank', 'Inter-American Development Bank' and 'Supra-National.' *Industrials* is comprised of 'Gas and Transmission', 'Utility and Electricity', 'Transport-Nonrail', 'Telephone' and 'Industrial.'

	Financials	Government agencies	Supra-Nationals	Industrials	Total
AAA	139	23	59	7	228
AA	99	33		13	145

for the rating classes in Table 1. The results for the industry classifications of Table 2 are similar and do not provide additional insights into the performance of the models. Table 3 summarises the specifications of the SC model. We use third degree *B*-splines for all bond categories, but the number and placement of the knot points differ. For the ratings GOVT, AAA, AA and A, we set the knots at 3 and 9 years, approximately corresponding to a segmentation into short-, medium- and long-term maturities. For the lower rating classes, BBB, BB and B, we use only one knot at 5 years, because the number of bonds in these classes is relatively small. The table also mentions the corrections to be applied to the AA-rated Frankfurt Interbank Offered Rates (FIBORs) in the calculation of the synthetic bonds' prices to take credit risk into account. The corrections are taken from Bloomberg and are inevitably approximations. For simplicity, we apply the same correction to all four synthetic bonds of a rating. The estimated curves are rather insensitive to small changes in these corrections, though completely omitting them is ill advised.

Table 4 summarises the term structure estimations for all 20 trading days of June 1998 by averaging our evaluation statistics for goodness of fit, smoothness and reliability. Fig. 2 graphically illustrates the SC estimation results for the first day of our data set, June 2nd, 1998, by depicting (a) spot interest rate curves with their 95% confidence intervals and spot spread curves, (b) residuals and (c) standard errors for the estimated discount curves. Note that we estimate discount curves and discount spread curves and that the corresponding spot curves and spot spread curves are subsequently calculated from these estimates.

The residual scatter plots in Fig. 2b show that goodness of fit decreases with the credit quality of the bonds. The market prices of government bonds are reasonably approximated by the theoretical DCF prices, since all absolute deviations, except one, are less than 0.25 basis points (bps). The scatters for the bond categories AAA, AA, A and BBB are more dispersed, but the absolute errors are

Table 3
Model specifications for single-curve and multi-curve models
Degree and *knots* refer to the specification of the *B*-splines. *FIBOR correction* is the spread (in basis points) applied to Frankfurt Interbank Offered Rates (FIBORs) in calculating the prices of synthetic bonds to take credit risk into account.

	Single-curve		Multi-curve		FIBOR correction
	Degree	Knots	Degree	Knots	
GOVT	3	3,9	3	3,9	−20
AAA	3	3,9	2	9	−10
AA	3	3,9	2	9	0
A	3	3,9	2	9	+10
BBB	3	5	2	5	+20
BB	3	5	2	5	+30
B	3	5	2	5	+40

Table 4
Summary statistics of single-curve (SC) and multi-curve (MC) estimates, averaged over all 20 trading days of June 1998
Goodness of fit is calculated as the Root Mean Squared Error (RMSE) of the residuals. *Smoothness* is measured as (10^8 times) the integral of the square of the second derivative, and is calculated for the spot curve and the spot spread curve. *Reliability* is (10^4 times) the standard error of the estimated spot curve, evaluated at maturities 2, 5 and 10 years.

		Goodness of fit	Smoothness		Reliability		
			Spot	Spread	2 *y*	5 *y*	10 *y*
GOVT	SC	0.09	4.5		1.5	1.7	4.6
AAA	SC	0.32	2.4	2.5	3.2	3.3	6.2
	MC	0.33	2.2	0.1	2.0	2.6	5.4
AA	SC	0.28	9.1	3.6	3.5	3.9	9.3
	MC	0.29	4.5	0.7	2.2	2.9	7.8
A	SC	0.32	23	7.5	6.2	6.5	19
	MC	0.32	4.6	0.9	3.7	4.7	14
BBB	SC	0.47	80	61	23	25	
	MC	0.52	6.8	1.5	17	19	
BB	SC	3.5	1049	1076	177	143	186
	MC	3.8	31	131	36	30	39
B	SC	5.3	689	594	209	383	
	MC	5.3	829	788	87	110	

still smaller than 100 bps. Fitted prices for BB- and B-rated bonds are the least accurate with the largest residuals being about 1000 bps. Similar conclusions are drawn from the RMSE statistic in Table 4, which has the lowest value for GOVT bonds and generally increases for lower rated bonds. Apparently, the bonds in lower rating categories are increasingly more heterogeneous, which can be attributed to relative differences in perceived default probabilities and recovery rates of the issuers. Liquidity differences between the bonds may also contribute to the larger dispersion of pricing errors.

The extent to which market prices can be fitted accurately, has consequences for the reliability of the estimated curves. For the GOVT, AAA, AA and A curves, the confidence bounds in Fig. 2a almost coincide with the spot curve, whereas for the BBB, BB and B curves the upper and lower bounds are distinctly observable in the graphs. Therefore, the reliability decreases with credit worthiness. The plotted standard errors in Fig. 2c and the reliability statistics in Table 4 also show that the estimation error increases in maturity segments that contain a small number of bonds.

Most interesting for our purposes are the credit spread curves that can be obtained by subtracting the estimated corporate spot curve from the estimated government spot curve. All spot spread curves of June 2nd, 1998 in Fig. 2a have an unrealistically twisting shape, because they have alternately positively and negatively sloped segments. Especially the spread curves for investment-grade

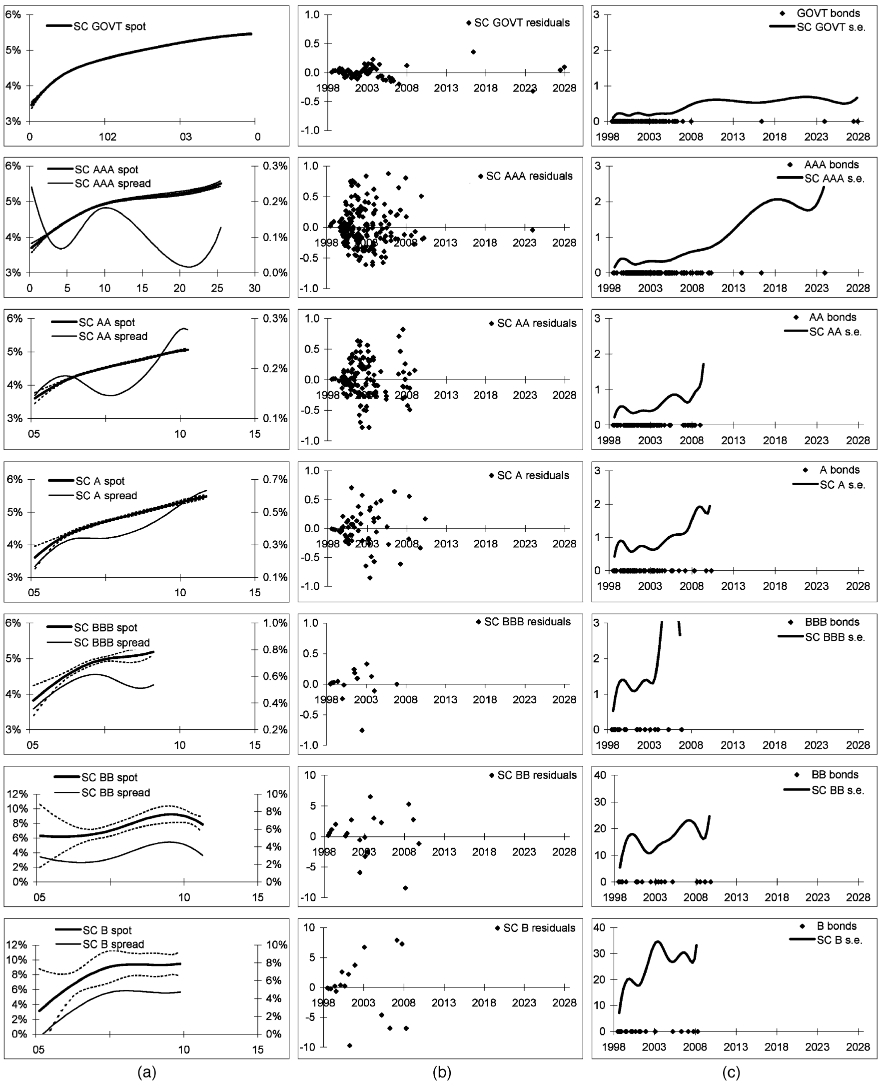

Fig. 2. Single-curve (SC) estimates for June 2nd 1998. Graphs (a) show estimated spot curves (with 95% confidence intervals) and spot spread curves; graphs (b) contain residuals; graphs (c) display (1000 times) the standard errors (s.e.) of estimated discount curves.

bonds display twisty behaviour. This can be explained from the relatively small magnitude of these spreads compared to interest rate levels: spot interest rates for classes AAA, AA, A and BBB are approximately 3.5–5.5%, whereas spot spreads range from 0.1% to 0.6%. Therefore, small deviations in either the government or the corporate spot curve can result in substantial irregularities in the spread; in other words, any dissimilarity in the curves' curvatures implies twists in the spread curve. These fluctuations are indeed observed in the graphs. The twisting spread curves are not in line with the smooth spread curves that are predicted by the theoretical bond price models of Merton (1974) and Longstaff and Schwartz (1995). The curves also contradict the empirical research by Helwege and Turner (1999), who found statistically significant evidence for increasing credit spreads for both investment-grade and speculative-grade issuers. Usage of these spread curves, e.g. for Value at Risk calculations or credit derivatives pricing, is likely to result in erroneous outcomes.

5.3. Multi-curve results

In the MC model we explicitly and parsimoniously model the spread curve, so that we hypothesise to obtain smooth spread curves. Since the model for a corporate bond category now only has to focus on the discount spread relative to the default-free government discount curve, together with the use of a combined data set of government and corporate bonds, we expect the parameters to be estimated more reliably. A consequence of the joint parameter estimation is that the parameters of the government spline model change somewhat compared to the SC estimates. Therefore, the government curve obtained from the MC model differs somewhat from the SC government curve. The changes are very small, though, and here we are only interested in the corporate curves.

Table 3 shows the specification of the degree and knots for the MC model. For the government discount curve, we use exactly the same settings as in the SC model, but for all corporate discount spread curves we lower the degree and for some curves also the number of knots. Rating categories AAA, AA and A use quadratic splines with one knot at 9 years, while the BBB, BB and B discount spread curves are modelled as quadratic splines with one knot at 5 years. Later we show how these specifications are found, but first we show that they indeed yield the hypothesised favourable properties of the MC model over its SC competitor. Note that we use the same corrections to FIBOR to value the synthetic zero-coupon bonds as in the SC estimations, so that any bias that may be caused by using incorrect values is equal for the SC and MC models.

Fig. 3 provides a graphical representation of the results of June 2nd, 1998 and Table 4 again summarises the main characteristics by averaging the goodness of fit, smoothness and reliability statistics over all 20 trading days of June 1998.

The graphs in Fig. 3a contain the estimated spot spread curves from the MC model and, for comparison, also the SC spot spreads from Fig. 2a. For all bond categories, there is a major improvement in the smoothness of the spread curves. Compared to the fluctuating SC spread curves, the MC spread curves are smooth, mostly increasing functions of time to maturity. For the longest maturities, however, all spread curves, except for ratings AA and A, have a negatively sloped

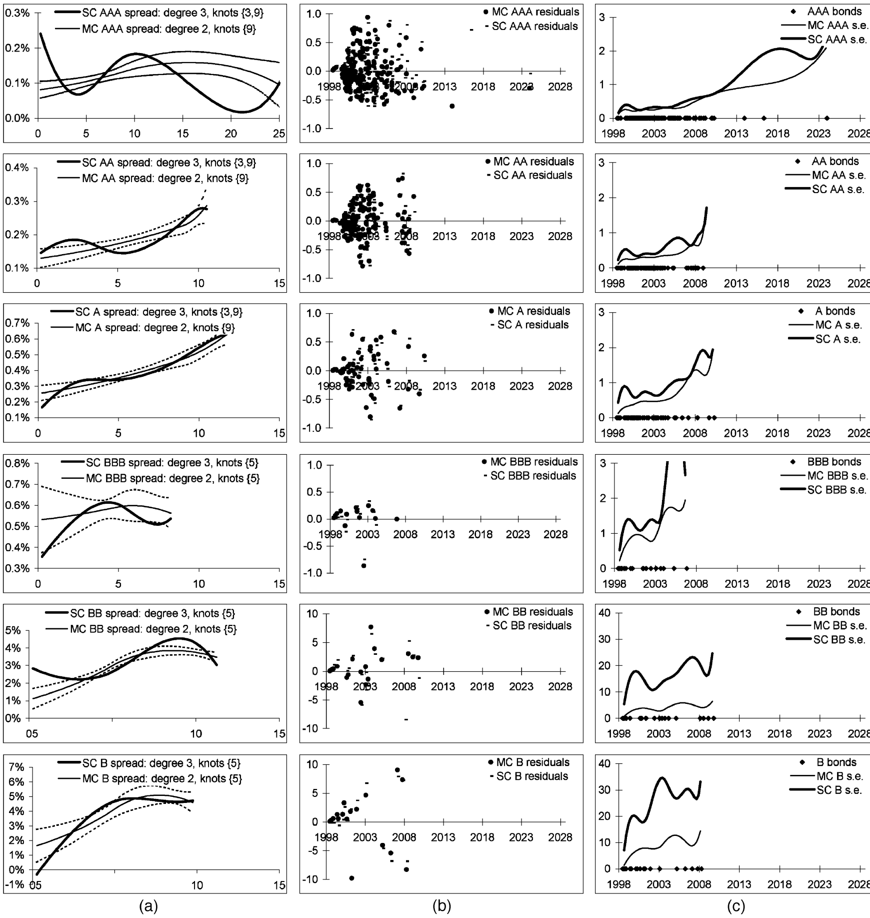

Fig. 3. Single-curve (SC) and multi-curve (MC) estimates for June 2nd, 1998. Graphs (a) show spot spread curves (with 95% confidence intervals for MC spreads); graphs (b) contain residuals; graphs (c) display (1000 times) the standard errors (s.e.) of estimated discount curves.

segment; this behaviour is likely to be caused by a sampling bias as discussed in Helwege and Turner (1999). Table 4 adds quantitative evidence to these graphical insights: the smoothness statistic is decimated for all rating categories, except B. The improvements range from a factor 5 for category A to 40 for BBB-rated bonds. Corporate spot curves, calculated as the sum of the estimated government and spread curves, also become smoother, again with category B as the only exception. Note that the smoothness statistics for the spot spread curves are now much smaller than for the spot curves, confirming our prior belief in Section 3 that spreads have a less complicated shape.

The second advantage of the MC model over independently estimated SC models resides in the increased reliability of the estimated corporate term structures.[5] Fig. 3c and Table 4 show that for all ratings and all maturities the standard errors of MC curves are lower than those of SC curves. The reliability for the BB and B categories improves most. For segments with a small number of bonds, such as the interval 2010–2020 for category AAA and 2004–2006 for AA, the standard errors become much smaller. This can be attributed to the combination of data sets, which improves the density of the distribution of maturity dates.

Since the total number of parameters in the MC model is smaller than in the corresponding SC models, the goodness of fit may be expected to decrease. This is also ‘predicted’ by the familiar trade-off between flexibility and smoothness. However, the price that we have to pay for the improvements in smoothness and reliability is rather modest, because the RMSE statistics in Table 4 are approximately equal for the SC and MC model for all rating categories. The scatter plots in Fig. 2b also show about the same dispersion of the residuals and do not reveal any significant biases. Apparently, the focus on the credit spread and the joint estimation with the default-free government curve, offset the negative effects of the imposed parsimonious structure.

5.4. Robustness

In this section we show that the shape of the spread curve in the MC model is relatively robust to the precise specification of the model, whereas the SC model is more sensitive to model misspecification. Along the way, we illustrate that the favourable properties of the MC model cannot be obtained from a parsimoniously specified corporate SC model or from a richly specified MC model. We show these results for the AAA-, AA- and A-curves, since for these curves we can consider more alternative degree and knot settings.

As starting point for the specification of the discount spread curves of AAA, AA and A in the MC model, we choose the specification of their discount curves in the SC model, i.e. a cubic spline with knot scheme {3, 9}. Subsequently, we reduce the flexibility of the spline model by lowering the degree or the number of knots (or both). Including the initial combination, this yields 6 alternative specifications: degree 3 and knots {3,9}; 3,{3}; 3,{9}; 2,{3,9}; 2,{3} and 2,{9}. The specification for the government curve is unchanged as a cubic spline with knots {3, 9}. Fig. 4 shows estimated spot spread curves for the three ratings using MC models with the six degree–knots combinations. For comparison, the figure also

[5] Note that we cannot judge the change in reliability of the *spread* curves, because we are unable to calculate standard errors of estimated spreads for the SC case. With SC models, the parameters of the government and corporate term structure models are estimated independently, so that we do not get the required covariances.

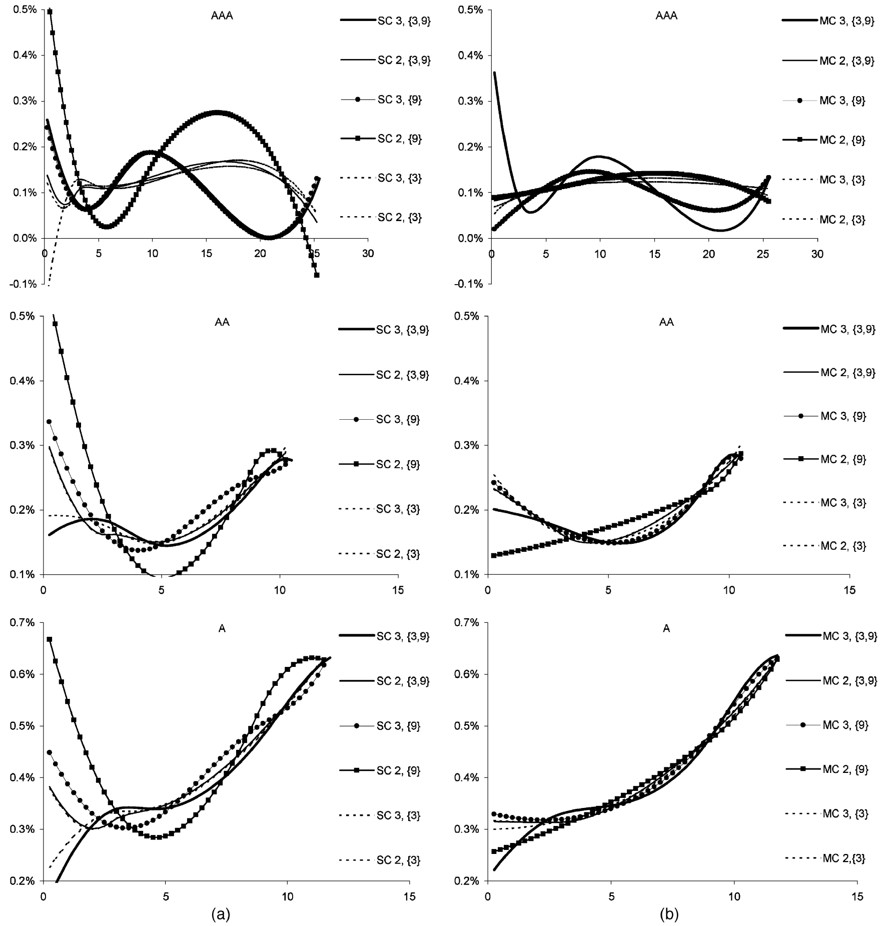

Fig. 4. Spot spread curves estimates for June 2nd, 1998 for ratings AAA, AA and A, estimated with single-curve (SC) models [graphs (a)] and multi-curve (MC) models [graphs (b)] with different degree and knot settings.

contains spot spread curves that are calculated from independently estimated SC models, where the specification of the GOVT discount curve is unchanged and the AAA-, AA- and A-discount curves are specified according to one of the six degree–knots combinations.

The bold lines in Fig. 4 depict the estimated SC and MC spreads for the first combination. The shape of the MC spreads are just as twisty as those of the SC models. It is therefore clear that an MC model with the same number of parameters as an SC model, does not yield the described favourable results of a parsimoniously specified MC model. We can smooth the spread curve by decreasing the spline degree and reducing the set of knots. We apply this idea for both the corporate discount curves in the SC model and the corporate discount spread curves in the MC model. The graphs in Fig. 4 reveal that the shapes of the SC spot spread curves calculated with these new settings vary considerably. This shows that estimating a parsimoniously specified SC model does not yield smooth and intuitively shaped spread curves. The spread curves obtained from the MC model, however, stiffen and are not as sensitive to different degree–knot settings. Particularly for AA and A, there is much less variation in shape and curvature of the curves. This robustness to model specification is a distinct advantage of the MC model over its SC competitor.

5.5. Spread curve specification

To choose between the competing parsimonious MC models of the previous section, we apply the Curve Similarity Test of Section 4. The most richly specified MC models, i.e. with the highest degree and the largest number of knots, are used as base case. These models are restricted by lowering the degree and/or reducing the number of knots. The null hypothesis in each CST claims that the spot spread curve of a restricted model is equal to the curve of the largest model if evaluated in a prespecified maturity vector. Section 4 discussed several criteria on the dimension and the spacing of this maturity vector; one choice that satisfies these criteria is to set the maturities at the spline knots and at maturities exactly between these knots. For rating AAA, for example, the knots of the largest MC model are set at {0, 3, 9, 25} (including the end points), so that we specify the maturities of the CST as {1.5, 3, 6, 9, 17}. Table 5 shows the maturity vectors for the other

Table 5
Results of Curve Similarity Tests applied to spot spread curves obtained from multi-curve models with different degree and knot settings
Maturities are the maturities in which the test statistic is evaluated; H_1 shows the degree and knots of the most richly specified multi-curve model that is used as alternative hypothesis; H_0 lists the restricted models and the average p-values of the test over all trading days of June 1998.

	Maturities	H_1	H_0					
			2, { 3,9}	3, {9}	2, {9}	3, {3}	2, {3}	2, {5}
AAA	1.5, 3, 6, 9, 17	3, {3,9}	0.00	0.16	0.00	0.00	0.00	
AA	1.5, 3, 6, 9, 10	3, {3,9}	0.92	1.00	0.09	0.76	0.93	
A	1.5, 3, 6, 9, 10.5	3, {3,9}	0.37	0.87	0.07	0.85	0.49	
BBB	2.5, 5, 6.5	3, {5}						0.07
BB	2.5, 5, 8	3, {5}						0.04
B	2.5, 5, 7.5	3, {5}						0.50

rating classes,[6] as well as the settings of the most richly specified MC model and of the restricted MC models.

For each trading day of June 1998, we calculate the CST statistic and confront the test statistics with critical values from a χ^2-distribibition to obtain 20 p-values for each rating and model specification. The averages of these p-values are reported in Table 5. We first consider the results of AA, a case where the CST is indeed helpful in the specification process. All alternative degree–knot combinations, except degree 2 with knot {9}, yield average p-values well in excess of any reasonable confidence level. This means that the spot spread curves corresponding to these combinations are—on average—not significantly different from the most richly specified MC model. In other words, these combinations result in spreads that are just as twisting as those from the SC model of Section 5.2. The model with degree 2 and one knot at 9, on the other hand, delivers spread curves that are on the edge of the rejection/non-rejection interval with an average p-value of 0.09. In the sense of Section 4, this model strikes the optimal balance between goodness of fit and smoothness. A similar conclusion may be drawn for ratings A, BBB and B, whose p-values range from 0.04 to 0.07.

For AAA-rated bonds, all models, except 3,{9}, are significantly different from the largest MC model. Taking into account the very twisting nature of the spread curve in Fig. 4b, this is hardly surprising. In this case, we feel that visual inspection of the estimated spread curves has to take prevalence over the use of the CST statistic. Thus, any of the four *rejected* degree–knots combinations may be used and for consistency with AA and A, we choose the model with degree 2 and one knot at 9. Likewise, for rating class B we use the same settings as for BBB and BB.

6. Conclusions

We present a new framework for the joint estimation of term structures and credit spreads. By decomposing a corporate term structure in a default-free curve and a credit spread curve, we can use a parsimoniously specified model for the spread curve and take the default-free part from the government curve. Both the government and the spread curve are modelled as B-splines and their parameters are jointly estimated from a combined data set.

We use a data set of liquid, German mark denominated bonds, with Standard and Poor's ratings ranging from AAA to B. For comparison, we first independently estimate separate *single-curve* term structure models for each of the rating

[6] We also conducted the test with other maturity vectors, where all maturities were either decreased or increased by 0.25 years. Since the results (which are available on request from the authors) were about the same, the CST statistic is fairly robust to the precise choice of the maturity vector.

classes. Spread curves that are obtained by subtracting the estimated government and corporate curves have unrealistically twisting shapes, and we argue that this can be attributed to the relatively small magnitude of spreads compared to interest rates. Therefore, any dissimilarity in the curvatures of the government and corporate curves implies twists in the spread curve.

Next we apply the proposed multi-curve model that explicitly and parsimoniously models the spread curve and jointly estimates it with the government curve. We illustrate that the model yields smooth spread curves that are more in line with the theoretical bond price models. Because the parameters are estimated from a combined data set of government and corporate bonds, the reliability of estimated term structures improves considerably. This effect is strongest for those maturity segments of a term structure that contain only a small number of bonds for that particular rating class. The ability of the multi-curve model to accurately fit market prices of bonds is hardly affected, in spite of the smaller number of parameters. We find that the favourable results of the model can be attributed to both the joint and the parsimonious modelling of the spread curves. To determine the optimal settings of the spline model for the spread curve, we use a newly developed test statistic that allows us to compare spot spread curves that are calculated from competing multi-curve models.

The new term structure estimation framework is valuable for each model that requires accurately estimated term structures for different credit risk classes. Examples are Value at Risk calculations for bond portfolios, pricing models for corporate bonds, models that value credit derivatives and models that asses credit risk in derivative products.

Acknowledgements

Part of the research was conducted while the first author was affiliated with the Quantitative Research Department of the Robeco Group. The authors thank Ton Vorst, the editor, the associate editor, and two anonymous referees for valuable suggestions and critical remarks; Michel van der Spek and Chios Slijkhuis for research assistance, Kees van den Berg and Peter Ferket for useful comments on earlier versions of the paper, and Rabobank International for providing the data. Financial support by the Erasmus Center for Financial Research (ECFR) is much appreciated. Views expressed in the paper are the authors' own and do not necessarily reflect those of Rabobank International or Robeco Group.

Appendix A. *B*-splines

Below we give a brief description of the construction of a basis of *B*-splines; see Powell (1981) for a more elaborate discussion and Steeley (1991) for an application of *B*-splines to term structure estimation.

Given $n+1$ knots $\tau_0 < \tau_1 < \dots < \tau_n$, the kth degree B-spline basis function is defined as:

$$B_s^k(t) = \sum_{l=s}^{s+k+1} \left[\prod_{h=s,h\neq l}^{s+k+1} \frac{1}{\tau_h - \tau_l} \right] \max(t-\tau_l,0)^k,$$

where subscript s denotes that this B-spline is only non-zero if t is in $[\tau_s, \tau_{s+k+1}]$. Powell (1981, p. 234) provided an efficient recurrence relation to evaluate the spline functions,

$$B_s^k(t) = \frac{(t-\tau_s)B_s^{k-1}(t) + (\tau_{s+k+1}-t)B_{s+1}^{k-1}(t)}{\tau_{s+k+1}-\tau_s},$$

with start conditions,

$$B_j^1(t) = \begin{cases} \dfrac{\tau_{i+1}-t}{(\tau_{i+1}-\tau_{i-1})(\tau_{i+1}-\tau_i)} & \text{if } j=i-1 \\ \dfrac{t-\tau_i}{(\tau_{i+1}-\tau_i)(\tau_{i-2}-\tau_i)} & \text{if } j=i \\ 0 & \text{if } j\neq i-1, j\neq i \end{cases}.$$

To construct a basis, we need $n+k$ linearly independent B-splines. Because a kth order B-spline is only non-zero in $k+1$ subintervals, within the interval $[\tau_0, \tau_n]$ only $n-k$ B-splines are defined. To construct a basis of $n+k$ functions, another $(n+k)-(n-k)=2k$ splines are required. A convenient way of choosing them so that they are also B-splines is to introduce extra knots $\{\tau_i;\ i=-k, -k+1, \dots, -1\}$ and $\{\tau_i;\ i=n+1, n+2, \dots, n+k\}$ outside the interval $[\tau_0, \tau_n]$. Commonly, these auxiliary knots are set as $\{\tau_i = \tau_0 + i(\tau_1-\tau_0);\ i=-k, -k+1, \dots, -1\}$ and $\{\tau_i = \tau_n + (i-n)(\tau_n - \tau_{n-1});\ i=n+1, n+2, \dots, n+k\}$. Then we construct a basis of $n+k$ B-splines consisting of $\{B_s^k;\ s=-k, -k+1, \dots, n-1\}$.

Appendix B. Variances and covariances

With spline estimation it is straightforward to calculate the variance of an estimated discount factor for a specific maturity t. Since the estimate $\hat{D}_c(t)$ of the discount factor for maturity t is a linear combination of the parameter estimates, we can apply the result that the variance of a linear combination $\mathbf{a}'\boldsymbol{\xi}$ of a vector of random variables $\boldsymbol{\xi}$ with covariance matrix $\mathrm{cov}(\boldsymbol{\xi})$ is given by the quadratic form:

$$\mathrm{var}(\mathbf{a}'\boldsymbol{\xi}) = \mathbf{a}'\mathrm{cov}(\boldsymbol{\xi})\mathbf{a}.$$

Hence,

$$\mathrm{var}\left(\hat{D}_1(t)\right) = \mathbf{g}_1(t)'\mathrm{cov}\left(\hat{\boldsymbol{\beta}}_1\right)\mathbf{g}_1(t),$$

and

$$\mathrm{var}\left(\hat{D}_c(t)\right) = \left(\mathbf{g}_1(t),\mathbf{g}_c(t)\right)'\mathrm{cov}\left(\hat{\boldsymbol{\beta}}_1,\hat{\boldsymbol{\beta}}_c\right)\left(\mathbf{g}_1(t),\mathbf{g}_c(t)\right), \quad c=2,3,\ldots,C.$$

Similarly, the variance of an estimated discount spread $\hat{s}_c(t) = \hat{D}_c(t) - \hat{D}_1(t)$ equals:

$$\mathrm{var}\left(\hat{s}_c(t)\right) = \mathbf{g}_c(t)'\mathrm{cov}\left(\hat{\boldsymbol{\beta}}_c\right)\mathbf{g}_c(t), \quad c=2,3,\ldots,C.$$

Because the spot curves and spot spread curves depend non-linearly on the estimated parameters, their variances are not straightforwardly computed. A useful approximation is given by the delta method (Greene, 2000, p. 118): the variance of a function $\phi(\cdot)$ of $\boldsymbol{\xi}$ is approximated by the quadratic form:

$$\mathrm{var}\left(\phi(\boldsymbol{\xi})\right) \approx \left(\frac{\partial\phi(\boldsymbol{\xi})}{\partial\boldsymbol{\xi}}\right)\mathrm{cov}(\boldsymbol{\xi})\left(\frac{\partial\phi(\boldsymbol{\xi})}{\partial\boldsymbol{\xi}}\right)'. \tag{B.1}$$

The estimated spot rate $\hat{r}_c(t)$ is a function of the estimated discount factor $\hat{D}_c(t)$,

$$\hat{r}_c(t) = \phi_1\left(\hat{D}_c(t)\right) = -\frac{\ln\hat{D}_c(t)}{t},$$

so that its variance is approximately,

$$\mathrm{var}\left(\hat{r}_c(t)\right) \approx \left(\frac{-1}{t\hat{D}_c(t)}\right)\mathrm{var}\left(\hat{D}_c(t)\right)\left(\frac{-1}{t\hat{D}_c(t)}\right) = \frac{\mathrm{var}\left(\hat{D}_c(t)\right)}{\left(t\hat{D}_c(t)\right)^2}.$$

To construct the variance for an estimated spot spread curve,

$$\hat{s}_{r,c}(t) = \hat{r}_c(t) - \hat{r}_1(t) = -\frac{\ln\left(\hat{d}(t) + \hat{s}_c(t)\right)}{t} + \frac{\ln\hat{d}(t)}{t}, \tag{B.2}$$

we write it explicitly as a function of the estimated parameters $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_c$,

$$\hat{\mathbf{s}}_{r,c}(t) = \phi_2\left(\hat{\boldsymbol{\beta}}_1,\hat{\boldsymbol{\beta}}_c\right) = -\frac{\ln\left(\mathbf{g}_1(t)'\hat{\boldsymbol{\beta}}_1 + \mathbf{g}_c(t)'\hat{\boldsymbol{\beta}}_c\right)}{t} + \frac{\ln\left(\mathbf{g}_1(t)'\hat{\boldsymbol{\beta}}_1\right)}{t}.$$

Using Eq. (B.1), the variance is then approximated as:

$$\mathrm{var}\left(\hat{s}_{r,c}(t)\right) \approx \left(\frac{\partial\hat{s}_{r,c}(t)}{\partial\hat{\boldsymbol{\beta}}_1},\frac{\partial\hat{s}_{r,c}(t)}{\partial\hat{\boldsymbol{\beta}}_c}\right)\mathrm{cov}\left(\hat{\boldsymbol{\beta}}_1,\hat{\boldsymbol{\beta}}_c\right)\left(\frac{\partial\hat{s}_{r,c}(t)}{\partial\hat{\boldsymbol{\beta}}_1},\frac{\partial\hat{s}_{r,c}(t)}{\partial\hat{\boldsymbol{\beta}}_c}\right)', \tag{B.3}$$

where

$$\frac{\partial \hat{s}_{r,c}(t)}{\partial \hat{\boldsymbol{\beta}}_1} = \frac{-\mathbf{g}_1(t)}{t\left(\mathbf{g}_1(t)'\hat{\boldsymbol{\beta}}_1 + \mathbf{g}_c(t)'\hat{\boldsymbol{\beta}}_c\right)} + \frac{\mathbf{g}_1(t)}{t\mathbf{g}_1(t)'\hat{\boldsymbol{\beta}}_1} \tag{B.4a}$$

and

$$\frac{\partial \hat{s}_{r,c}(t)}{\partial \hat{\boldsymbol{\beta}}_c} = \frac{-\mathbf{g}_c(t)}{t\left(\mathbf{g}_1(t)'\hat{\boldsymbol{\beta}}_1 + \mathbf{g}_c(t)'\hat{\boldsymbol{\beta}}_c\right)}. \tag{B.4b}$$

The idea of a variance of a spot spread at a single maturity can be extended to a full covariance matrix of a vector containing spreads for several maturities. Let $\hat{\mathbf{s}}_{r,c}(\mathbf{t})$ be a vector function of the maturity vector $\mathbf{t} = (t_1, \dots, t_q)'$,where the i-th element is given by the spot spread function (Eq. (B.2)) evaluated in maturity t_i. Then we can apply a vector version of Eq. (B.1) to obtain a result similar to Eq. (B.3):

$$\Sigma_c^1(\mathbf{t}) = \operatorname{cov}\left(\hat{\mathbf{s}}_{r,c}(\mathbf{t})\right) \approx \left(\frac{\partial \hat{\mathbf{s}}_{r,c}(\mathbf{t})}{\partial\left(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_c\right)}\right) \operatorname{cov}\left(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_c\right) \left(\frac{\partial \hat{\mathbf{s}}_{r,c}(\mathbf{t})}{\partial\left(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_c\right)}\right)',$$

where $\partial \hat{\mathbf{s}}_{r,c}(\mathbf{t}) / \partial(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_c)$ is a $q \times (n_1 + k_1 + n_c + k_c)$ matrix with the i-th row equal to Eqs. (B.4a) and (B.4b) evaluated in t_i.

References

Adams, K.J., van Deventer, D.R., 1994. Fitting yield curves and forward rate curves with maximum smoothness. Journal of Fixed Income 4 (1), 52–62.

Anderson, N., Breedon, F., Deacon, M., Derry, A., Murphy, G., 1996. Estimating and Interpreting the Yield Curve. Whiley, Chicester.

Beim, D.O., 1992. Term structure and the non-cash value in bonds, Working Paper 92-40. Graduate School of Business, Columbia University.

Black, F., Derman, E., Toy, W., 1990. A one-factor model of interest rates and its application to treasury bond options. Financial Analyst Journal 46 (1), 33–39.

Bliss, R., 1997. Testing term structure estimation methods. Advances in Futures and Options Research 9 (1), 197–231.

Brealey, R.A., Meyers, S.C., 1991. Principles of Corporate Finance. 4th edn. McGraw-Hill, New York.

Bühler, W., Uhrig-Homburg, M., Walter, U., Weber, T., 1999. An empirical comparison of forward- and spot-rate models for valuing interest-rate options. Journal of Finance 54 (1), 269–305.

Caouette, J.B., Altman, E.I., Narayanan, P., 1998. Managing Credit Risk: The Next Great Financial Challenge. Wiley, New York.

Duffee, G.R., 1996. On measuring credit risks of derivative instruments. Journal of Banking and Finance 20, 805–833.

Duffie, D., Singleton, K.J., 1999. Modeling term structures of defaultable bonds. Review of Financial Studies 12 (4), 687–720.

Greene, W.H., 2000. Econometric Analysis. 4th edn. Prentice Hall, New Jersey.

Heath, D., Jarrow, R.A., Morton, A., 1992. Bond pricing and the term structure of interest rates: a new methodology. Econometrica 60 (1), 77–105.

Helwege, J., Turner, C.M., 1999. The slope of the credit yield curve for speculative-grade issuers. Journal of Finance 54 (5), 1869–1884.

Hull, J., White, A., 1990. Pricing interest-rate-derivative securities. Review of Financial Studies 3 (4), 573–592.

Hull, J., White, A., 1995. The impact of default risk on the prices of options and other derivative securities. Journal of Banking and Finance 19, 299–322.

Jarrow, R.A., Turnbull, S.M., 1995. Pricing derivatives with credit risk. Journal of Finance 50 (1), 53–85.

Jarrow, R.A., Lando, D., Turnbull, S.M., 1997. A Markov model for the term structure of credit spreads. Review of Financial Studies 10 (2), 481–523.

Langetieg, T.C., Smoot, J.S., 1989. Estimation of the term structure of interest rates. Research in Financial Services 1, 181–222.

Litterman, R., Iben, T., 1991. Corporate bond valuation and the term structure of credit spreads. Journal of Portfolio Management, 52–64, Spring.

Litzenberger, R.H., Rolfo, R., 1984. An international study of tax effects on government bonds. Journal of Finance 39 (1), 1–22.

Longstaff, F.A., Schwartz, E.S., 1995. A simple approach to valuing risky fixed and floating rate debt. Journal of Finance 50 (3), 789–819.

McCulloch, J.H., 1971. Measuring the term structure of interest rates. Journal of Business 44, 19–31.

McCulloch, J.H., 1975. The tax-adjust yield curve. Journal of Finance 30 (3), 811–830.

Merton, R.C., 1974. On the pricing of corporate debt: the risk structure of interest rates. Journal of Finance 29 (2), 449–469.

Nelson, C.R., Siegel, A.F., 1987. Parsimonious modeling of yield curves. Journal of Business 60, 473–489.

Poirier, D.J., 1976. The Econometrics of Structural Change. North Holland Publishing.

Powell, M.J.D., 1981. Approximation Theory and Methods. Cambridge Univ. Press, Cambridge.

Saunders, A., 1999. Credit Risk Management: New Approaches to Value at Risk and Other Paradigms. Wiley, New York.

Schaefer, S.M., 1981. Measuring a tax-specific term structure of interest rates in the market for British government securities. Economic Journal 91, 415–438.

Shea, G.S., 1984. Pitfalls in smoothing interest rate term structure data: equilibrium models and spline approximations. Journal of Financial and Quantitative Analysis 19 (3), 253–269.

Shea, G.S., 1985. Interest rate term structure estimation with exponential splines: a note. Journal of Finance 40 (1), 319–325.

Steeley, J.M., 1991. Estimating the gilt-edged term structure: basis splines and confidence intervals. Journal of Business Finance and Accounting 18 (4), 513–529.

Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Research 5 (2), 177–188.

Vasicek, O.A., Fong, H.G., 1982. Term structure modeling using exponential splines. Journal of Finance 37 (2), 339–348.