

A Robust Approach to the Interest Rate Term Structure Estimation

Min Li¹ and Yan Yu²

¹California State University, Sacramento and ²University of Cincinnati

Abstract: This paper estimates the interest rate term structures of Treasury and individual corporate bonds using a robust criterion. The Treasury term structure is estimated with Bayesian regression splines based on nonlinear least absolute deviation. The number and locations of the knots in the regression splines are adaptively chosen using the reversible jump Markov chain Monte Carlo method. Due to the small sample size, the individual corporate term structure is estimated by adding a positive parametric credit spread to the estimated Treasury term structure using a Bayesian approach. We present a case study of U.S. Treasury STRIPS (Separate Trading of Registered Interest and Principal of Securities) and AT&T bonds from April 1994 to December 1996. Compared with several existing term structure estimation approaches, the proposed method is robust to outliers in our case study.

Key words: Coupon bonds, credit spreads, forward rate, regression splines, reversible jump Markov chain Monte Carlo, treasury bonds.

1. Introduction

It is well known that least squares estimates can be sensitive to outliers. Unfortunately, bond prices often exhibit heavy tails with possible outliers (Schwartz 1998; Jarrow, Ruppert, and Yu 2004). Schwartz (1998) uses a robust measure and finds that almost 10% of the US Treasury securities in the Fixed Income database are outliers. This paper considers term structure estimation of Treasury and corporate bonds using a robust approach. We extend the normal linear Bayesian regression splines model of Denison, Mallick, and Smith (1998) to the *nonlinear* term structure model based on the least absolute deviation criterion.

The term structure of interest rates of the bonds describes how interest rates evolve over time. The interest rate term structures of both government and corporate bonds have important applications in economics and finance, as described in detail in Li and Yu (2005). Obtaining accurate estimates of both Treasury and corporate term structures is essential.

Many methods have been developed for term structure estimation. For example, McCulloch (1971, 1975) and Shea (1984) use regression splines but with fixed knots to estimate the discount function. Chambers, Carleton, and Waldman (1984) fit non-linear least squares on the exponent of the discount function. Fisher, Nychka, and Zervos (1995), JRY (Jarrow, Ruppert, and Yu 2004) and Li and Yu (2005) present penalized (smoothing) splines approaches where the forward rate curve is smoothed with roughness penalty. In all these works, the least squares method is used and is not robust to outliers.

Schwartz (1998) estimates the term structure by modeling the forward interest rate with a piecewise constant curve using fixed knots, minimizing the usual sum of squared residuals. A subjective measure is then used to find and remove outliers in the estimation process. Many outliers are found. According to Schwartz (1998), there could be outliers in bond prices in the term structure data due to a variety of reasons: out-of-date bond ratings; large spreads in illiquid bonds; non-synchronous trades or data entry errors; use of the wrong interest rate function. However, the use of a piecewise constant curve in Schwartz (1998) for the interest rate gives large fitted bond pricing errors.

This paper presents a robust approach to term structure estimation by adopting the nonlinear least absolute deviation criterion. We model the interest rate with Bayesian regression splines. From our limited experience, this method is robust to outliers and is able to price bonds accurately. The number and locations of the knots are adaptively chosen by the reversible jump Markov chain Monte Carlo (MCMC) method of Green (1995). Most of previous term structure methods consider using regression splines or penalized (smoothing) splines with fixed knots. For example, in Shea (1984), the number and locations of the knots are specified subjectively and unrealistic term structure shape occurred. In JRY and Li and Yu (2005)'s penalizes spline approaches, the number and locations of the knots are fixed and smoothing is achieved through a roughness penalty.

With very few individual corporate bonds available, it is challenging to obtain both good term structure estimation and valid statistical inference. This paper estimates the term structures of individual corporate bonds by borrowing strength from (adding a credit spread to) the estimated Treasury term structure with a Bayesian approach as in Li and Yu (2005), but under Laplace errors. A credit spread is the excess return on a corporate bond over an equivalent Treasury bond. The knowledge that the credit spread is positive can be easily incorporated into our Bayesian model as prior information. Moreover, statistical inference is straightforward within this Bayesian framework.

We present a case study of U.S. Treasury STRIPS and AT&T coupon bonds on October 31, 1995 to illustrate the effectiveness of the proposed method in handling outliers. A rather extreme outlier is created and several competing

methods are compared with the proposed method. The proposed method is able to mitigate the effect of the outlier better than the other methods.

The remainder of the paper is organized as follows. We present a robust approach to estimating the Treasury term structure in Section 2. We then illustrate the Bayesian estimation of individual corporate term structure with informative priors in Section 3. Section 4 contains the case study. Some concluding remarks follow in Section 5.

2. A Robust Model for the Treasury Term Structure of Interest Rates

This section presents the robust procedure for Treasury term structure estimation. Section 2.1 outlines the term structure model. Section 2.2 presents the Bayesian estimation procedure using the reversible jump MCMC algorithm. The details of using the reversible jump MCMC method to choose the number and locations of the knots are presented in Section 2.3.

2.1 Term structure model

The term structure of interest rates of the bonds can be determined by any of the discount function $D(0, T)$, the yield curve $y(0, T)$, or the forward rate $f(0, T)$, where T is time to maturity from today (time 0). If we consider current time to be fixed at 0, the relationships among these three functions are:

$$D(T) = \exp\{-Ty(T)\} = \exp\left\{-\int_0^T f(s)ds\right\} \quad (2.1)$$

Once one of the three is obtained, the rest of them are determined. The discount function $D(T)$ represents today's price of a *zero-coupon bond* (paying no interest or principal until maturity) that pays one dollar at maturity date T . The yield curve $y(T) = \int_0^T f(s)ds/T$ is the average of the forward rates between today and the maturity date T . Intuitively, the forward rate $f(s)$ is the rate one can lock in today for future borrowing or lending at time s . It is a variable interest rate and gives the marginal return. To estimate the term structure, a spline may be used to model the discount function $D(T)$, the yield curve $y(T)$, or the forward rate $f(s)$. However, it is best to estimate the term structure of the forward rate, as argued in Fisher *et al.* (1995) and Li and Yu (2005).

The forward rates in term structure are not observable but implied by the price of the bonds. The forward rates and thus the term structure can only be estimated using a pricing model. The price of a risk free zero-coupon bond $P(T)$ on date 0 maturing on date T with a face value of \$1 is the same as the discount price $D(T)$ in our notation. The "empirical" forward rate (JRY) can be found by $f = \Delta\{-\log(P(T))\}/\Delta T$, where Δ is the differencing operator. It provides a

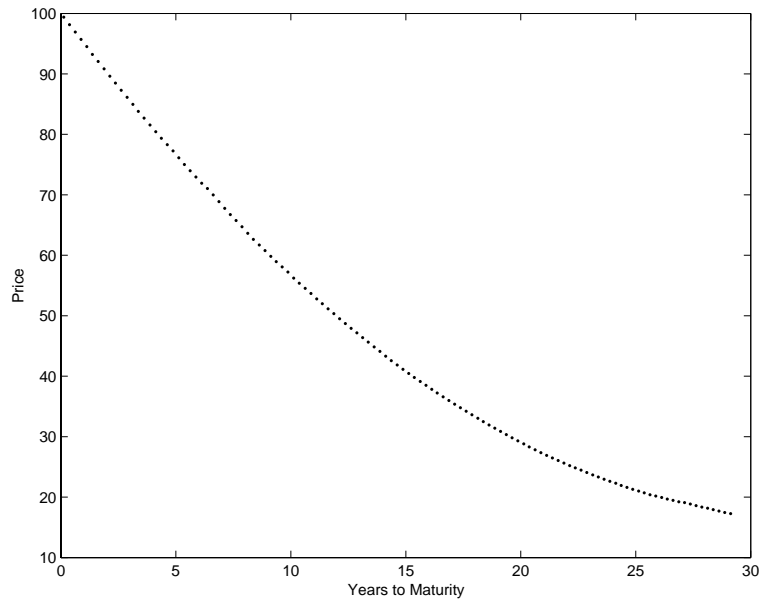


Figure 1: Price versus years to maturity for treasury STRIPS on October 31, 1995.

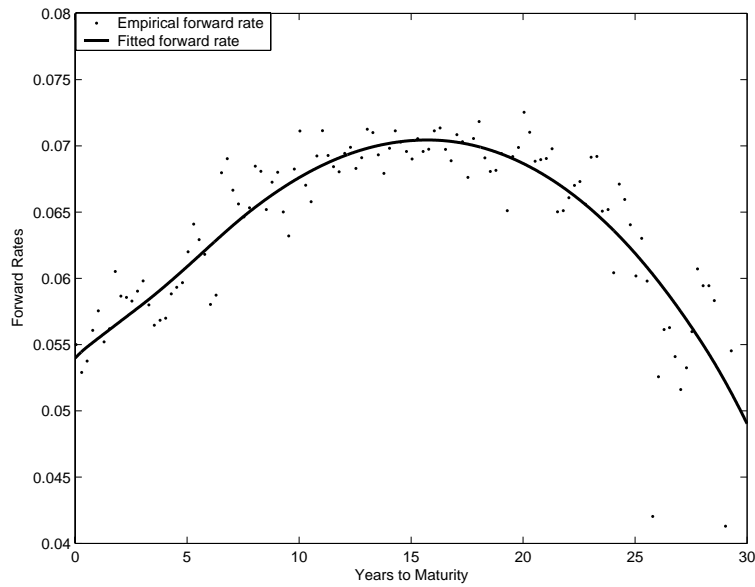


Figure 2: Empirical forward rates and the fitted forward rate curve (using the reversible jump MCMC and minimizing the sum of the absolute residuals) versus years to maturity for treasury STRIPS on October 31, 1995. Note that the empirical forward rates are not the true observed data (forward interest rates are not observable). The fitted forward rate curve is not obtained from fitting a model to the empirical forward rates but rather by fitting model (2.5) to the price as explained in Section 2.2.

rough estimate of the unobservable forward rate. Figure 1 and Figure 2 plot the price $P(T)$ and the empirical forward rates versus years to maturity T for U.S. Treasury STRIPS on October 31, 1995. Our objective is to estimate the forward rate based on the bond price.

Many bonds make regular coupon payments during the life of the bonds. Coupon bonds may be viewed as a collection of *zero coupon bonds* and the following notations are adopted. Each bond pays fixed coupons and principal $C_i(t_{ij})$ due on dates t_{ij} , where $j = 1, \dots, z_i$ and z_i is the total number of coupon and principal payments for the i -th bond. Therefore, t_{i1} is the first coupon payment date and t_{iz_i} is the maturity date of bond i . The model price for the coupon bond, μ_i , is related to the forward rate f through the discount function:

$$\mu_i = \sum_{j=1}^{z_i} C_i(t_{ij})D(t_{ij}) = \sum_{j=1}^{z_i} C_i(t_{ij}) \exp\left\{-\int_0^{t_{ij}} f(s)ds\right\}, \quad (2.2)$$

where the current time is assumed to be 0. Let $P(i), i = 1, \dots, n$, be the market price of the i -th bond on the current date (date 0). Given the observed price P_i of the i -th coupon bond, the statistical model becomes:

$$P_i = \mu_i + \epsilon_i \quad (2.3)$$

where ϵ_i is the error term.

2.2 A Bayesian regression spline model for the term structure of treasury bonds

The Treasury forward rate curve $f(s)$ is modeled with a regression spline $f(s) = \boldsymbol{\delta}'\mathbf{B}(s)$, where $\boldsymbol{\delta}$ is a vector of spline coefficients and $\mathbf{B}(s)$ is a vector of some basis functions. Let P_i be the i -th observed bond price and $\mu_i(\boldsymbol{\delta})$ be the model price from the regression spline model. For a fixed spline basis function $\mathbf{B}(s)$, the sum of the absolute deviations between and $\mu_i(\boldsymbol{\delta})$ is minimized:

$$\min_{\boldsymbol{\delta}} \sum_{i=1}^n |P_i - \mu_i(\boldsymbol{\delta})|.$$

The estimator obtained from this criterion has been shown by Koenker and Bassett (1978) to be resistant to outliers in the data.

For regression splines, the number and locations of the knots need to be selected. Many methods have been proposed: Friedman (1991), Smith and Kohn (1996), Denison *et al.* (1998, 2002), Lindstrom (1999), Zhou and Shen (2001), DiMatteo, Genovese, and Kass (2001), etc. Denison *et al.* (1998) apply the reversible jump MCMC (Green, 1995) scheme to select locations of the knots in

the linear regression model. After knot locations are selected using the reversible jump MCMC, the coefficient estimates are computed using the least squares method. All these smoothing methods are for linear models (linear in coefficients).

The term structure model (2.3) is *nonlinear* in the regression spline coefficients. We extend Bayesian regression splines of Denison *et al.* (1998) using nonlinear least absolute deviation. As in Denison *et al.* (2002), the number and locations of the knots are selected using the reversible jump MCMC method. To find the coefficient estimates for the nonlinear term structure function, we implement an iterative re-weighted least squares method, minimizing the sum of the absolute residuals. This can be regarded as a hybrid Bayesian and frequentist approach. A fully Bayesian procedure (e.g. DiMatteo *et al.* 2001) would be to impose priors on the coefficients. However, computation becomes too burdensome in the reversible jump MCMC implementation.

For the basis function in the regression spline model, we adopt the d -th degree truncated power basis:

$$f(s) = \boldsymbol{\delta}'\mathbf{B}(s) = \delta_0 + \delta_1 s + \cdots + \delta_d s^d + \sum_{k=1}^K \delta_{d+k} (s - t_k)_+^d, \quad (2.4)$$

where $\boldsymbol{\delta} = [\delta_0, \delta_1, \dots, \delta_{d+K}]'$, $\mathbf{B}(s) = [1, s, \dots, s^d, (s - t_1)_+^d, \dots, (s - t_K)_+^d]'$, $(s - t_k)_+ = \max(0, s - t_k)$, and $\{t_k\}_{k=1}^K$ are K spline knots. The number of knots K and locations of the knots t_i will be adaptively chosen. $\delta_0, \delta_1, \dots, \delta_d$ are the polynomial coefficients and $\delta_{d+1}, \dots, \delta_{d+K}$ are the spline coefficients. From our experience, $d = 2$ is sufficient for the data available. Power basis has the advantage of being simple and allows easily modeling of credit spreads for corporate term structure in the next section. Moreover, setting some polynomial coefficients to zero allows convenient modeling of sub-models. Other basis, such as B-spline (de Boor 1978), can also be used. However, truncated power basis can be easily transformed during computation to B-spline basis for numerical stability, as in many regression packages (Ruppert, Wand, and Carroll 2003, P. 71).

The model relating the observed market price P_i and the model price μ_i , defined in (2.2), for the i -th bond is: $P_i = \mu_i + \epsilon_i$. In matrix notations, the model we are working with is

$$\mathbf{P} = \boldsymbol{\mu} + \mathbf{e} \quad (2.5)$$

with

$$\begin{aligned} \mathbf{P} &= [P_1, P_2, \dots, P_n]' \\ \boldsymbol{\mu} &= [\mu_1, \mu_2, \dots, \mu_n]' \end{aligned}$$

$$\mu_i = \sum_{j=1}^{z_i} C_i(t_{ij}) \exp\left\{-\int_0^{t_{ij}} \delta' \mathbf{B}(s) ds\right\},$$

where t_{ij} is the years to maturity for the j -th coupon payment of the i -th bond and z_i is the total number of coupon and principal payments for the i -th bond. We integrate the forward rate:

$$\int_0^{t_{ij}} f(s) ds = \int_0^{t_{ij}} \delta' \mathbf{B}(s) ds = \delta' \int_0^{t_{ij}} \mathbf{B}(s) ds = \delta' \mathbf{B}^I(t_{ij}), \quad (2.6)$$

where

$$\begin{aligned} \mathbf{B}^I(t_{ij}) &= \int_0^{t_{ij}} \mathbf{B}(s) ds \\ &= \left[t_{ij}, \frac{t_{ij}^2}{2}, \dots, \frac{t_{ij}^{d+1}}{d+1} + \dots + \frac{(t_{ij} - t_1)_+^{d+1}}{d+1}, \dots, \frac{(t_{ij} - t_K)_+^{d+1}}{d+1} \right]. \end{aligned} \quad (2.7)$$

Then the model price for the i th coupon bond can be expressed as

$$\mu_i = \sum_{j=1}^{z_i} C_i(t_{ij}) \exp\{-\delta' \mathbf{B}^I(t_{ij})\}. \quad (2.8)$$

The parameters to be estimated are K, t_1, \dots, t_K , and σ^2 . The number K and locations of the knots t_1, \dots, t_K are selected by the reversible jump MCMC method and details are provided in the next section. Once the number of knots K and the knot locations t_1, \dots, t_K are selected, the coefficients $\delta_0, \delta_1, \dots, \delta_d, \delta_{d+1}, \dots, \delta_{d+K}$ can be determined from the data (knot locations) by applying a nonlinear iteratively re-weighted least squares method (see Seber and Wild 1989). This nonlinear iteratively re-weighted least squares method is outlined in the appendix. σ^2 is generated in a Gibbs step after the coefficients are estimated (the detail is provided in the next section). A number of simulation steps are carried out. To start, a few equal spaced knots (1, 2, or 3 knots equally spaced) are selected. The estimation steps are summarized in the algorithm at the end of Section 2.3. The iterative re-weighted least squares method automatically weighs large residuals (outliers) less and removes them in effect. This is in contrast with Schwartz (1998) who adopts a subjective measure to manually remove the outliers.

2.3 Reversible jump MCMC to select the number and locations of the knots

We use $\boldsymbol{\theta}$ to denote the parameter vector: $\boldsymbol{\theta} = (K, t_1, \dots, t_K, \sigma^2)$. Minimizing the sum of absolute residuals is equivalent to assuming the errors to be the Laplace distribution. The likelihood is

$$L_K(\mathbf{P} | \boldsymbol{\theta}) = \frac{1}{(\sqrt{2\pi}\sigma)^2} \exp \left\{ \frac{-1}{\sigma} \sum_{i=1}^n |P_i - \mu_i| \right\}. \quad (2.9)$$

We follow the general framework by Green (1995) and Denison *et al.* (2002): including a birth step (add a knot), a death step (delete a knot), and a movement step (move a knot) in the reversible jump MCMC algorithm. We model the forward rate curve with a spline as in (2.4):

$$f(s) = \delta_0 + \delta_1 s + \dots + \delta_d s^d + \sum_{k=1}^K \delta_{d+k} (s - t_k)_+^d.$$

Let K be the number of knots in the model with knots located at: (t_1, t_2, \dots, t_K) . Both K and knot locations (t_1, t_2, \dots, t_K) are random variables and need to be estimated. Green (1995) adopts the Poisson prior on K and the Poisson mean needs to be specified. A better choice is the discrete uniform distribution or the Geometric distribution (Smith and Kohn 1996, Hansen and Kooperberg 2002). We impose the discrete uniform distribution as the prior on K . This ensures that the number of knots is not influenced by subjective priors. Given the number of knots K , locations of the knots $\mathbf{t} = (t_1, t_2, \dots, t_K)$ are also assumed to be from a uniform distribution (see Denison *et al.* 2002). Each model with K knots should be equally likely to be selected and this probability is $K!(n_{knots} - K!)/n_{knots}!$, where n_{knots} is the total number of candidate knot locations. Any combination of K knots are distinct and selected with replacement. The ordering of these K knots does not matter in the reversible jump MCMC knots selection process. Thus, the prior probability for the locations of the knots becomes $p(\mathbf{t}) = [K!(n_{knots} - K!)/n_{knots}!] \times (1/k_{max} + 1)$ ($K \leq k_{max}$, where k_{max} is the maximum number of knots chosen to be in the model and n_{knots} is the total number of candidate knot locations. Candidate locations of knots are at data points. The birth, death, and move probabilities are chosen to be $b_K = d_K = m_K = 1/3$ ($K = 2, \dots, k_{max} - 1$) with $b_1 = m_1 = 1/3, b_{k_{max}} = d_1 = 0$, and $d_{k_{max}} = m_{k_{max}} = 1/2$. Any of the three birth, death, and move steps gives a candidate model.

The acceptance probability for the birth step is the product of the likelihood ratio, the prior ratio, the proposal ratio, and the Jacobian. The likelihood functions for both the current model and the candidate model are evaluated using

the parameter estimates of $\delta_0, \delta_1, \dots, \delta_d, \delta_{d+1}, \dots, \delta_{d+K}$, and σ^2 . The polynomial and spline coefficient estimates are computed using the iterative re-weighted algorithm outlined in the appendix. The variance can be treated as an auxiliary parameter. The inverted Gamma distribution $IG(a, b)$ is chosen to be the prior distribution for σ^2 . The conditional posterior distribution of σ^2 given the other parameters and data is proportional to

$$\sigma^{-2(a+n/2+1)} \exp\left(-\frac{1}{b\sigma^2} - \frac{1}{\sigma} \sum_{i=1}^n |P_i - \mu_i|\right). \quad (2.10)$$

Here a is chosen to be a small positive number and b is chosen to be a large positive number. Therefore, the prior distribution of σ^2 is the usual diffuse prior and has little impact on the posterior distribution. Since the conditional posterior distribution of is not in a closed form, a Metropolis step is used to generate its estimate.

Once all parameter estimates are computed, the likelihood ratio can be formed:

$$\frac{\text{Laplace}(\boldsymbol{\mu}_c, \sigma_c^2 I_n)}{\text{Laplace}(\boldsymbol{\mu}, \sigma_c^2 I_n)} \quad (2.11)$$

where the subscript c denotes the candidate model and $\text{Laplace}(\boldsymbol{\mu}_c, \hat{\sigma}_c^2 I_n)$ denotes the Laplace density function with corresponding mean and variance matrix. This ratio is the Bayes factor, approximately. However, DiMatteo *et al.* (2001) point out the likelihood ratio in the normal (least squares) model always favors the model with more parameters. This is also true for the least absolute deviation model here. DiMatteo *et al.* (2001) argue that the likelihood ratio based on the least squares methods in Denison *et al.* (1998) should include a penalty factor as in BIC. We also replace (2.11) with BIC (Schwarz 1978) to penalize the dimensionality of the model. According to Schwarz (1978), $-2 \log(\text{Bayes Factor})$ can be approximated by BIC:

$$BIC = -2 \log \left[\frac{\text{Laplace}(\boldsymbol{\mu}_c, \sigma_c^2 I_n)}{\text{Laplace}(\boldsymbol{\mu}, \sigma_c^2 I_n)} \right] - (\text{Param}_c - \text{Param}) \log n,$$

where the likelihood is evaluated at the maximum likelihood estimates, Param_c is the number of parameters in the candidate model, and Param is the number of parameters in the current model. Thus, the Bayes factor is approximately $\exp[-BIC/2]$. This relationship is used to approximate (2.11) and thus to penalize more complex models. To compute the acceptance ratio, the prior ratio and proposal ratio also need to be found. The prior ratio is $(K+1)/(n_{knots} - K)$ and the proposal ratio is $(d_{K+1}/b_K)[(n_{knots} - K)/(K+1)]$. The Jacobian equals

1 in that the model space is discrete. Thus, the acceptance probability for the birth step is

$$\alpha = \min\{1, \exp(-BIC/2) \times d_{K+1}/b\}. \quad (2.12)$$

Similarly, the acceptance probability for the death step is

$$\alpha = \min\{1, \exp(-BIC/1) \times b_{K-1}/d_K\}. \quad (2.13)$$

Each collection of knot locations represents a possible model and our task is to choose the correct model. One approach to find the correct model is to choose the mode, i.e., the knot locations that give the maximum likelihood for the model. An alternative approach is to adopt the view of Bayesian model averaging and find the weighted average of the estimated values for f based on a large number of simulations from the predictive distributions. It is well-known that given the predictive squared-error loss function for the model response, the Bayes estimator is the expected response under the predictive distribution. This is equivalent to finding the average of all models considered. Thus, given N reversible jump MCMC samples, the posterior mean of f can be approximated by

$$\frac{1}{N} \sum_{i=1}^N E\{f \mid \mathbf{B}, \mathbf{P}, \boldsymbol{\theta}^{(i)}\}. \quad (2.14)$$

The algorithm is outlined below.

1. Initially, select k_0 knot locations on data points.
2. Compute the spline coefficient estimates of δ 's using the nonlinear iterative re-weighted least squares method described in the appendix.
3. Generate a uniform (0,1) random number v :
 - (i) If $v < b_k$, perform the birth step. Add a knot by randomly choosing a knot from the candidates. Perform step 2 and compute the acceptance probability (2.12). Accept the birth step with the acceptance probability.
 - (ii) If $b_k < v \leq b_k + d_k$ perform the death step. Randomly choose a knot from the current model and delete it. Perform step 2 and compute the acceptance probability (2.13). Accept the deletion with the acceptance probability.
 - (iii) If $v > b_k + d_k$, perform the move step. Relocate a knot and compute the acceptance probability (the penalized likelihood ratio), $\exp(-BIC/2)$. The move is accepted with the acceptance probability.
4. Repeat the steps 2 and 3 above until the mean absolute deviation ($MAD = (1/n) \sum_{i=1}^n |P_i - \mu_i|$) converges.

5. Find the forward rate curve either using the mode or using the model averaging estimate (2.14).

3. Bayesian Estimation of the Individual Corporate Term Structure

In this section, we estimate the term structures of individual corporate bonds. The method here is similar to Li and Yu (2005) with the following differences: (i) the errors are assumed to follow the Laplace distribution, not normal; (ii) the type of credit spreads (constant, linear, or quadratic) is selected based on the posterior odds ratio instead of the posterior interval. Only a brief description is provided here and more details can be found in Li and Yu (2005).

Given the small sample size of individual corporate bonds traded on any given day, we “borrow strength” from the estimated Treasury term structure and estimate the credit spread between the two: $f_c = f_{Tr} + \text{spread}$, where f_c is the corporate term structures and $f_{Tr} = \tilde{\boldsymbol{\delta}}' \mathbf{B}^I(t_{ij})$ is the estimated Treasury term structure from Section 2. A credit spread represents the excess return from a corporate bond over the return from an equivalent Treasury bond. There is one Treasury term structure curve with a large sample size but many corporate bonds with very small sample sizes each. It is natural to estimate the Treasury term structure once and then estimate different individual corporate term structures by finding the spreads. For a constant spread, $f_c = f_{Tr} + \beta_0 = \boldsymbol{\delta}'_c \mathbf{B}(s)$, where $\boldsymbol{\delta}_c = [\delta_0 + \beta_0, \delta_1, \delta_2, \dots, \delta_{d+K}]$. The linear and quadratic spreads can be similarly expressed. The sum of the absolute residuals from the corporate bond prices is minimized. This is equivalent to having both the errors and the likelihood function follow the Laplace distribution. For computation speed, the estimated Treasury term structure used here is the model with the knot locations that give the maximum likelihood, not the moving average estimate.

We first consider the estimation of a constant spread. As in Li and Yu (2005), the spread is assumed to be distributed uniformly between 0 and a positive constant g prior to estimation: $[\beta_0] = (1/g)I_{(0,g)}(\beta_0) \propto I_{(0,g)}(\beta_0)$. The posterior distribution for the constant spread is

$$[\beta_0 | \mathbf{P}, \sigma^2] \propto [\beta_0] [\mathbf{P} | \beta_0, \sigma^2] \propto I_{(0,g)} \exp\left(-\frac{1}{\sigma} \sum_{i=1}^n |P_i - \mu_i|\right),$$

given the errors (and thus the likelihood) are Laplace. This is not of any known distribution. Adopting the usual inverted Gamma distribution $IG(a, b)$ (small a and large b) as the prior for the error variance, the corresponding posterior distribution is again not a standard distribution:

$$[\sigma^2 | \mathbf{P}, \beta_0, a, b] \propto [\sigma^2] [\mathbf{P} | \beta_0, \sigma^2]$$

$$\propto \sigma^{-2(a+n/2+1)} \exp\left(-\frac{1}{b\sigma^2} - \frac{1}{\sigma} \sum_{i=1}^n |P_i - \mu_i|\right). \quad (3.1)$$

The Metropolis algorithm is applied for both posteriors to generate sample draws. The linear $(\beta_0 + \beta_1 s)$ and quadratic $(\beta_0 + \beta_1 s^2)$ spreads can be estimated similarly.

For inference, the posterior interval, a $100(1 - \alpha)\%$ central interval, can be obtained easily from MCMC sample draws. Choosing the type of spread (constant, linear, or quadratic) is the same as choosing the correct model. The posterior odds can be used. For example, to select between a constant spread or a linear spread, the posterior odds ratio

$$\frac{[\beta_0|\mathbf{P}]}{[\beta_0, \beta_1|\mathbf{P}]} = \frac{[\beta_0][\sigma_{Const}^2][\mathbf{P}|\beta_0, \sigma_{Const}^2]/[\mathbf{P}]}{[\beta_0, \beta_1][\sigma_{Lin}][\mathbf{P}|\beta_0, \beta_1, \sigma_{Lin}^2]/[\mathbf{P}]} \approx \frac{[\mathbf{P}|\beta_0 \sigma_{Const}^2]}{[\mathbf{P}|\beta_0, \beta_1, \sigma_{Lin}^2]}, \quad (3.2)$$

can be computed to choose the model. Here σ_{Const}^2 and σ_{Lin}^2 are error variances from the constant and linear spread models respectively. The priors on the model parameters are flat and only the likelihood ratio needs to be computed. However, the posterior odds ratio is only an approximation from the MCMC sample draws.

Table 1: AT&T Bonds on October 31, 1995 from the fixed Income Database. The coupon listed is the semi-annual dollar coupon payment.

Current date	Issue date	Maturity date	Coupon	Price
Oct 31, 1995	Jan 14, 1992	Jan 15, 2002	7.125	106.279
Oct 31, 1995	March 24, 1994	April, 2004	6.750	102.406
Oct 31, 1995	June 1, 1994	June 1, 2006	7.500	110.197
Oct 31, 1995	Feb 28, 1995	March 1, 2007	7.750	110.175
Oct 31, 1995	May 12, 1995	May 15, 2005	7.000	106.583

4. A Case Study

A case study is presented in this section to illustrate the proposed method. The data we use are the end-of-month U.S. Treasury STRIPS and AT&T Bonds from April 30, 1994 to December 31, 1995. The U.S. Treasury STRIPS are *zero coupon bonds* that are synthesized from the coupon and principal payments of Treasury bonds. These data are obtained from the University of Houston Fixed Income Database. The database contains over 28,000 instruments, including publicly traded non-convertible debt with principal value no less than one million dollars. The bond data that make up the Lehman Brothers Bond Indices are reported with month-end flat prices, accrued interest, coupon, yields, current

date, issue date, maturity date, S&P and Moody's ratings, and option-like features. The market price of a corporate bond equals the quoted flat price plus the accrued interest. The time period the database covers is from January 1973 to March 1997 and there have been no updates after March 1997.

There were five AT&T bonds traded on October 31, 1995. These five bonds have semi-annual coupons with different maturities and have no embedded options feature, e.g., the right to prepay, for which our model does not apply. Table 1 shows current dates, issue dates, maturity dates, coupons, and market prices of these five AT&T bonds. Treasury STRIPS data have the same format with no coupons.

For estimation, the time-to-maturity and coupon payment times need to be converted to the same unit scale. The converted current dates, issue dates, maturity dates, and first coupon payment dates in terms of years from current dates in Table 1 are listed in Table 2.

Table 2: AT&T Bonds on October 31, 1995. Dates and first coupon payment time are converted to units of one year based on actual/365 day count. The current date is set to time 0. The coupon listed is the semi-annual dollar coupon payment.

Date (year)	Issue (yr)	Maturity (yr)	First coupon	Coupon	Price
0	-3.7945	6.2082	0.2082	7.125	106.279
0	-1.6055	8.4164	0.4164	76.750	102.406
0	-1.4164	10.5836	0.0849	7.500	110.197
0	-6712	11.3315	0.3315	7.750	110.175
0	-30.4712	9.5370	0.0411	7.000	106.583

There are usually few individual corporate bonds available in a given month. In fact on average, there were only 4.3 bonds available per month during the period of April 1994 to December 1995. It is difficult to obtain a meaningful estimate of the individual corporate term structure based on these few observations. On the other hand, the average number of U.S. Treasury STRIPS per month during the same period was 117, ranging from 115 to 120. Thus, "borrowing strength" from other sources such as U.S. Treasury STRIPS becomes necessary in estimating the term structure of individual corporate bonds. Below the Treasury term structure is estimated first and then the AT&T term structure is estimated by adding a credit spread to the former.

We focus on one STRIPS data set on October 31, 1995 first and use it to illustrate the robustness of the proposed method by introducing a noticeable outlier to the STRIPS prices. The raw prices range from 15.29 to 99.79 while the years to maturity range from 0.04 years to 29.29 years. One price maturing in 27.79 years is \$16.6. We change this price to \$30 to make it an outlier and use this

rather extreme outlier to demonstrate the effectiveness of the proposed method. Comparison is made with several competing methods: the least squares (normal, not Laplace errors) and penalized splines approaches. The least squares (LS) method is implemented with the same knot selection scheme as in Section 2. The frequentist penalized splines (Jarrow *et al.* 2004) method minimizes penalized sum of squared errors and selects the smoothing parameter with EBBS (Empirical Bias Bandwidth Selection) of Ruppert (1997). The Bayesian penalized splines (BPS) method (Li and Yu 2005) assumes the errors to be normal. This is equivalent to minimizing penalized sum of squared errors but the smoothing parameter is automatically obtained as a by-product (a ratio of posterior variances) in estimation.

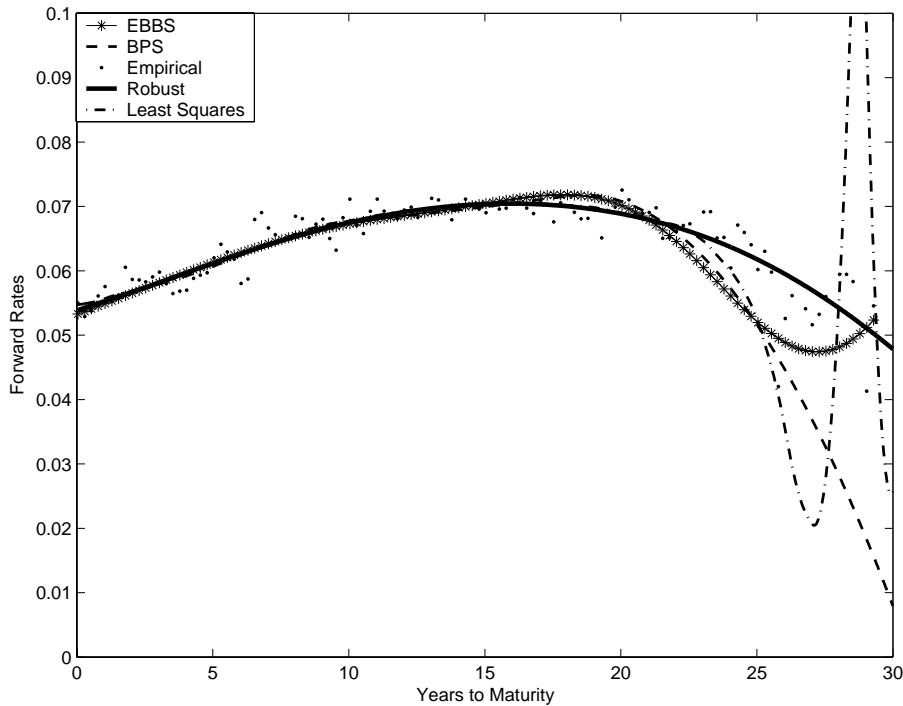


Figure 3: The fitted forward rate curves for treasury STRIPS on October 31, 1995 using the proposed robust approach and other competing methods. Note that the empirical forward rates are not the true observed data (forward interest rates are not observable), but approximated (see Section 2.1). The STRIPS price outlier at 27.79 years to maturity causes the empirical rate at that point to swing dramatically, very different from the others. The fitted forward rate curve is not obtained from fitting a model to the empirical forward rates but rather by fitting model (2.5) to the price as explained in Section 2.2. The graph has to be scaled and the outlier is not shown so that the fitted curves can be clearly shown.

The fitted curves from all methods are plotted in Figure 3. The outlier causes a complete breakdown of the least squares method at the outlier, 27.79 years to maturity (the graph has to be scaled and the outlier is not shown so that the fitted curves can be clearly shown). Both frequentist (EBBS) and Bayesian (BPS) penalized splines methods perform better but are still affected by the outlier. On the other hand, the robust method proposed is little affected by the outlier. We compute the mean absolute deviation (with the outlier included) $MAD = \sum_{i=1}^n |P_i - \mu_i|/n$, where μ_i is the model price. The MAD for the proposed robust method is considerably smaller than the others: 0.16 versus 0.24 (LS), 0.26 (BPS), and 0.26 (EBBS). In Schwartz (1998), outliers are identified using a subjective measure and are then removed. The method proposed here does not require us to identify or remove the outliers. Instead, it naturally weighs outliers less through weighing the residuals in the iterative re-weighted least squares method and the fitted curve is little affected by the outlier, as shown in the previous example.

The proposed method is also used to fit the STRIPS data on October 31, 1995 (without the outlier introduced earlier) and the fitted forward rate curve is displayed in Figure 2 in Section 2.1. The results are based on the moving average estimate (2.14).

Given this estimated Treasury term structure of interest rates, we apply the method outlined in Section 3 to estimate the AT&T term structure on October 31, 1995 using the five AT&T bonds available on that date. A spread is added to the estimated Treasury term structure and is estimated. 2,000 simulation draws (with another 2,000 draws for the burn-in period) from the Metropolis method are used to find the spread estimates and construct the posterior intervals. We observe that 2,000 burn-in draws are more than enough before the Markov chains mix adequately. First, a constant spread is added to the estimated Treasury term structure and is found to be 0.0043 with a 95% posterior interval (0.0041, 0.0044). The proposal distribution is normal and only the samples that satisfy the constraint are used.

A linear spread, $\beta_0 + \beta_1 s$, is then considered. The two parameter estimates are found to be 0.0036 and 0.00017. The 95% posterior intervals for both parameters are $(-0.0003, 0.0063)$ and $(-0.0005, 0.0011)$. The posterior odds ratio for a constant spread compared to a linear spread is found to be 9.97 and the constant spread model is favored. At maturity time 0, the risk of default by AT&T should be negligible and the credit spread should be due to liquidity risk, not credit risk. The name "credit spread" is a misnomer since it can be due to other factors in addition to credit differences (see Marshall 2000). AT&T bonds are not as liquid as Treasury bonds and they may not be sold immediately if necessity arises. A premium in the interest rate is needed to compensate for this liquidity risk and find an immediate buyer. The intercept of the credit spread can be interpreted

as the liquidity risk and it does exist based on the posterior odds ratio. However, the posterior odds ratio merely means that a constant spread is more likely and it does not mean the spread is indeed constant. A sample size of only 5 may not be sufficient to detect a difference from a constant spread. A constant spread model can serve as a good first approximation for AT&T and many other corporate bonds with a minimal number of bonds available. For individual corporate bonds with larger sample size, non-constant spreads may very well be found.

We also consider a quadratic spread, $\beta_0 + \beta_1 s^2$, and find the posterior odds ratio to be around 1. The quadratic spread model is not favored and a simpler constant spread is clearly a better choice. The 95% posterior intervals for both parameters are (0.0032, 0.005) and (-0.00003, 0.00004). The case of including both the linear and the quadratic spread is not considered for this AT&T data set. It is not desirable to estimate three parameters with only five data points. However, for other individual corporate bonds with a larger sample size, including both the linear and the quadratic spreads may be considered.

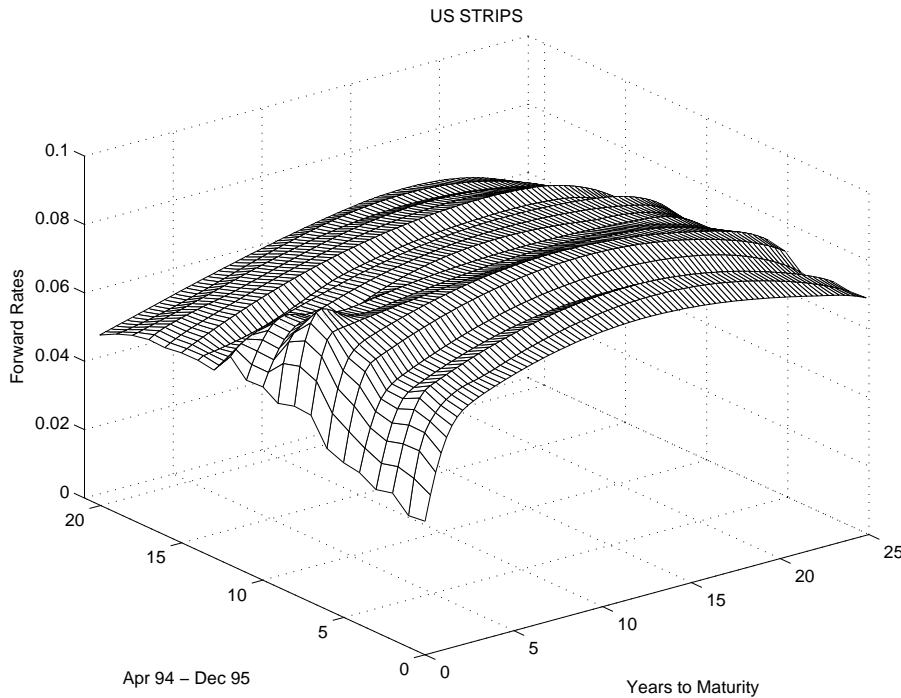


Figure 4: Fitted forward rate curves for US STRIPS over the 21-month period of April 1994 to December 1995.

We have focused on the term structure of one data set of October 31, 1995. However, modeling the evolution (over time) of the term structures of both Treasury and individual corporate bonds is an important problem in pricing interest

rate derivatives, among other applications in finance (Jarrow 2002). To study this evolution, the method described in Section 2 is applied 21 times independently to the 21 end-of-month U.S. Treasury STRIPS data sets from April 1994 to December 1995. Figure 4 displays the 21 fitted forward rate curves. The STRIPS data are sparse after 25 years to maturity and only the forward rate curves up to 25 years to maturity are plotted. The figure shows the evolution of the end-of-month forward rates over 21 months as well as the forward rates based on years to maturity from 0 to 25 years. The forward rate can be viewed as a function of time at fixed maturity.

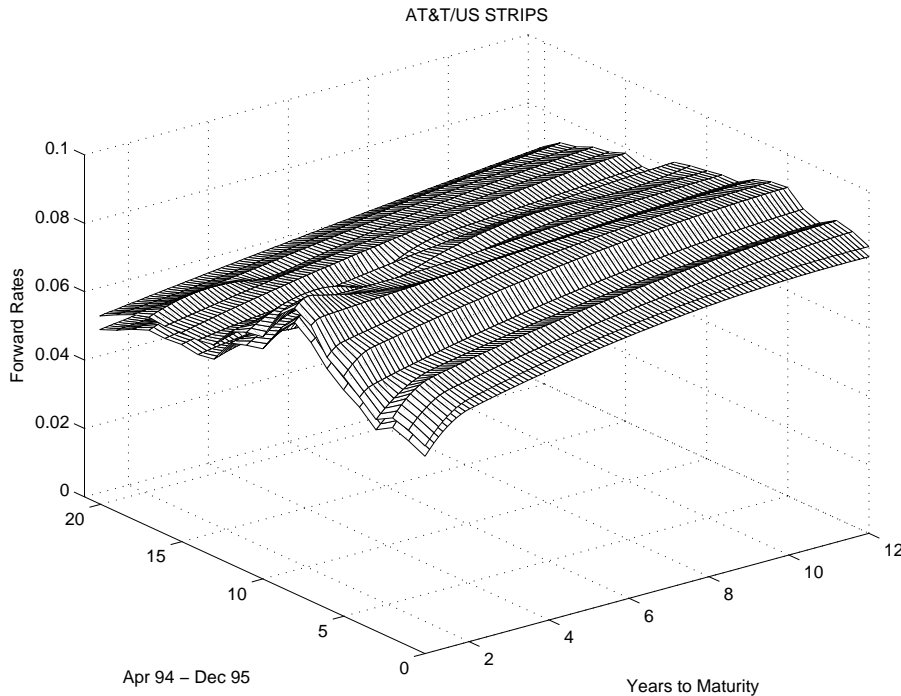


Figure 5: Fitted forward rates curves for AT&T (upper sheet) and US STRIPS (lower sheet) over the 21-month period of April 1994 to December 1995.

Given the 21 estimated STRIPS forward rate curves, we estimate the term structures of the 21 end-of-month AT&T bond data sets from April 1994 to December 1995 by finding the spreads above the estimated STRIPS term structure. A constant spread is favored according to the posterior odds ratios. Figure 5 plots the fitted forward rates for U.S. STRIPS and AT&T bonds over the 21-month period of April 1994 to December 1995. The upper sheet contains 21 AT&T forward rate curves with constant spreads and the bottom sheet contains the 21 STRIPS forward rate curves. The range of years to maturity for these

21 end-of-month AT&T data sets is from 0.9 to 12 years and is used in the figure. Extrapolation beyond this range is not desirable. We notice that at a fixed maturity, the forward rate curves (as a function of time) are rough. This is not surprising in that interest rates move randomly and abruptly. Thus, smoothing both maturity and time with a bivariate function is not desired.

5. Conclusions

In this paper, a robust approach is proposed to estimate term structures of both Treasury and individual corporate bonds. For the Treasury term structure, the forward rates are modeled with Bayesian regression splines and the number and locations of the spline knots are adaptively chosen with the reversible jump MCMC algorithm. The proposed approach is found to be robust against outliers in bond prices. A Bayesian method is then used to estimate the term structure of individual corporate bonds with very small sample sizes by borrowing strength from the estimated Treasury term structure.

We need to point out here that though the proposed robust approach shows promises in the case study, further research is necessary on possible error heteroscedacity and correlation. We hope to address these in another paper. Credit risk derivatives pricing using the robustly estimated corporate bond term structure may be very much worth further exploring.

Acknowledgement

The authors thank the editor for his constructive remarks. The first author thanks the 2003 summer fellowship support by the Office of Research and Sponsored Projects, California State University, Sacramento.

Appendix. A Nonlinear Iteratively Re-weighted Least Squares Method¹

1. Choose some starting values for and find the residual $e_i = P_i - \mu_i$.
2. Form a diagonal matrix \mathbf{U} with $\sqrt{e_i}$ as the diagonal elements.
3. Let \mathbf{P} be a vector of P_i 's. Find $\mathbf{Z} = \mathbf{P}\mathbf{U}^{-1}$.
4. Let $\boldsymbol{\mu}$ be a vector of μ_i 's and find $\mathbf{A} = \boldsymbol{\mu}\mathbf{U}^{-1}$.
5. Find

$$\mu_i^{[1]} = \left. \frac{\partial P_i(\boldsymbol{\delta})}{\partial \boldsymbol{\delta}'} \right|_{\boldsymbol{\delta}} = - \sum_{j=1}^{z_i} C_i(t_{ij}) \exp\{\boldsymbol{\delta}'\mathbf{B}^I(t_{ij})\} \mathbf{B}^I(t_{ij})$$

¹See Seber and Wild (1989) for computing the coefficients Given the Selected number and locations of the knots using the reversible jump MCMC method.

and use it to form the $(n \times (1 + d + K))$ matrix $\boldsymbol{\mu}^{[1]}$.

6. At each iteration, find $\mathbf{Ad}^{(iter)} = [\boldsymbol{\mu}^{(1)}(\boldsymbol{\delta}^{[iter]})]\mathbf{U}^{-1}$. Update $\boldsymbol{\delta}$ with a Gauss-Newton step

$$\boldsymbol{\delta}^{iter+1} = \boldsymbol{\delta}^{iter} + \left(\mathbf{Ad}^{iterprime} \mathbf{Ad}^{(iter)} \right)^{-1} [\mathbf{Z} - \mathbf{A}(\boldsymbol{\delta}^{(iter)})].$$

References

- Chambers, D. R., Carleton, W. T., and Waldman, D. W. (1984). A new approach to estimation of the term structure of interest rates. *Journal of Financial and Quantitative Analysis* **19**, 233-252.
- De Boor, C. (1978). *A Practical Guide to Splines*. Springer Verlag.
- Denison, D. G. T., Holmes, C., Mallick, B. K., and Smith, A. F. M. (2002). *Bayesian Methods for Nonlinear Classification and Regression*. Wiley.
- Denison, D. G. T., Mallick, B. K., and Smith, A. F. M. (1998). Automatic Bayesian curve fitting. *Journal of the Royal Statistical Society, Series B* **60**, 330-350.
- DiMatteo, I., Genovese, C. R., and Kass, R. E. (2001). Bayesian curve-fitting with free-knot splines. *Biometrika* **88**, 1055-1071.
- Fisher, M., Nychka, D., and Zervos, D. (1995). Fitting the term structure of interest rates with smoothing splines. Working Paper, Finance and Economics Discussion Series, Federal Reserve Board.
- Friedman, J. H. (1991). Multivariate adaptive regression splines (with discussion). *Annals of Statistics* **19**, 1-141.
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**, 711-732.
- Hansen, M. and Kooperberg, C. (2002). Spline adaptation in extended linear models. *Statistical Science* **17**, 2-51.
- Jarrow, R., Ruppert, D., and Yu, Y. (2004). Estimating the term structure of corporate debt with a semiparametric penalized spline model. *Journal of the American Statistical Association* **99**, 57-66.
- Koenker, R. and Bassett, G. Jr. (1978). Regression quantiles. *Econometrica* **46**, 33-50.
- Li, M. and Yu, Y. (2005). Estimating the interest rate term structures of treasury and corporate debt with Bayesian penalized splines. *Journal of Data Science* **3**, 223-240.
- Lindstrom, M. J. (1999). Penalized estimation of free-knot splines. *Journal of Computational and Graphical Statistics* **8**, 333-352.
- Marshall J. F. (2000). *Dictionary of Financial Engineering*. Wiley.

- 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**, 811-830.
- Ruppert, D. (1997). Empirical-bias bandwidths for local polynomial nonparametric regression and density estimation. *Journal of the American Statistical Association* **92**, 1049-1062.
- Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric Regression*. Cambridge University Press.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461-464.
- Schwartz, T. (1998). Estimating the term structure of corporate debt. *Review of Derivatives Research* **2**, 193-230.
- Seber, G. A. F. and Wild, C. J. (1989). *Nonlinear Regression*. Wiley.
- Shea, G. (1984). Pitfalls in smoothing interest rate term structure data: Equilibrium models and spline approximations. *Journal of Financial and Quantitative Studies* **19**, 253-269.
- Smith, M. and Kohn, R. (1996). Nonparametric regression using Bayesian variable selection. *Journal of Econometrics* **75**, 317-343.
- Zhou, S. and Shen, X. (2001). Spatially adaptive regression splines and accurate knot selection schemes. *Journal of the American Statistical Association* **96**, 247-259.

Received November 22, 2004; accepted January 26, 2005.

Min Li
College of Business Administration
California State University, Sacramento
6000 J Street
Sacramento, CA 95819-6088, USA
limin@csus.edu

Yan Yu
Department of Quantitative Analysis and Operations Management
PO Box 210130
Cincinnati, OH 45221-0130, USA
Yan.Yu@uc.edu