European Journal of Operational Research 204 (2010) 343–354

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

A relaxed cutting plane algorithm for solving the Vasicek-type forward interest rate model

Homing Chen [a], Cheng-Feng Hu [b,*]

[a] Department of Money and Banking, National ChengChi University, Taipei 115, Taiwan
[b] Department of Industrial Engineering and Management, I-Shou University, Ta-Hsu, Kaohsiung 84001, Taiwan

ARTICLE INFO

Article history:
Received 20 June 2008
Accepted 12 September 2009
Available online 17 September 2009

Keywords:
Semi-infinite programming
Forward interest rate

ABSTRACT

This work considers the solution of the Vasicek-type forward interest rate model. A deterministic process is adopted to model the random behavior of interest rate variation as a deterministic perturbation. It shows that the solution of the Vasicek-type forward interest rate model can be obtained by solving a nonlinear semi-infinite programming problem. A relaxed cutting plane algorithm is then proposed for solving the resulting optimization problem. The features of the proposed method are tested using a set of real data and compared with some commonly used spline fitting methods.

1. Introduction

The interest rate model plays a central role in the theory of modern economics and finance. In past studies interest rate models described by stochastic processes were widely used. It is often assumed that a small number of interest rates are sufficient statistics for the stochastic movement of the entire term structure. An enormous amount of work has been directed towards modeling and estimation of the short term interest rate dynamics. Some single-factor models [7,9,37] have been proposed and widely used in practice because of their tractability and their ability to fit reasonably well the dynamics of short term interest rates. Econometric estimation of these models has also been intensively studied in the literature [7,11,12,1,25]. Recently, Kortanek and Medvedev [19] introduced a deterministic process to model the random behavior of interest rate variation as a deterministic perturbation which was later investigated by Staffa [30], and Tichatschke et al. [34]. Moreover, McCulloch [22], Carleton and Cooper [5], Schaefer [28], Vasicek and Fong [38], Chambers et al. [6], Nelson and Siegel [24], Steeley [32] and Pham [26] used curve fitting techniques with the observed government coupon bond prices to estimate the pure discount bond yield curve. As it is well known, the use of piece-wise quadratic functions to approximate the yield curve may generate a non-smooth forward yield curve [21], and the use of polynomial functions to fit the entire yield curve may lead to unacceptable yield patterns [20]. Brown and Dybvig [3], Brown and Schaefer [4], de Munnik and Schotman [23] and Sercu et al. [29] used economic functions such as the Vasicek [37] and the Cox et al. [9] term structure models to fit the market yield curve. Although these functions can provide economic explanations, they fail to provide a rich variety of shapes to fit the versatile market yield curve [21]. Inspired and motivated by the recent research, this work considers the Vasicek-type forward interest rate model, which contains a one-dimensional source of randomness affecting bond prices (i.e. one-dimensional Brownian motion). For a fixed current time t, the Vasicek-type forward interest rate model considers the differentiation of the forward interest rate with respect to the time to maturity τ, which can be described as follows:

$$df(\tau|t) = (\alpha(t) + \beta(t)f(\tau|t))d\tau + dB(\tau|t), \quad \forall \tau \geqslant t,\ f(t|t) = r(t), \tag{1}$$

where f is the instantaneous forward interest rate, $B(\tau|t)$ is the one-dimensional standard Brownian motion with $B(t|t) = 0$, the increment $B(\tau|t) - B(s|t)$ is normally distributed with mean zero and variance $\tau - s$, for $t \leqslant s \leqslant \tau$ [18], the coefficients $\alpha(t)$, $\beta(t)$, and the spot rate $r(t)$ satisfy the following conditions:

$$0 < \underline{\alpha}(t) \leqslant \alpha(t) \leqslant \bar{\alpha}(t), \quad \underline{\beta}(t) \leqslant \beta(t) \leqslant \bar{\beta}(t) < 0, \quad 0 < \underline{r}(t) \leqslant r(t) \leqslant \bar{r}(t),$$

with the pre-assigned bounds $\underline{\alpha}(t),\ \bar{\alpha}(t),\ \underline{\beta}(t),\ \bar{\beta}(t),\ \underline{r}(t),\ \bar{r}(t)$. While the bounds on the short rate are imposed for reasons of compatibility with the impulse perturbation formulation below, it should be noted that a negative short rate would imply arbitrage in a world in which one

* Corresponding author. Tel.: +886 7 6577711x5530; fax: +886 7 6578536.
E-mail address: chu1@isu.edu.tw (C.-F. Hu).

0377-2217/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2009.09.012

can always hold cash [33], and the mean-reverting property (which is implied by the bounds on β) implies that extremely high (or low) values of r are extremely unlikely. We assume the spot rate is bounded less than infinity and greater than zero in this work.

The main feature of the Vasicek-type forward interest rate model (1) is the instantaneous trend of the process to revert to its long run mean value. The parameter $\beta(t)$ determines a speed of the adjustment and should be negative to ensure convergence [19]. The Vasicek-type model has also been extended in subsequent research. The work of Dothan [10], Courtadon [8], Cox et al. [9] and Stapleton and Subrahmanyam [31] can also be placed in this category.

To solve the Vasicek-type forward interest rate model (1), the concept of deterministic perturbations is adopted to deal with the random behavior of interest rate variations. It is shown that the solution of the Vasicek-type forward interest rate model (1) can be obtained by solving a nonlinear semi-infinite programming problem. A relaxed cutting plane algorithm is proposed for solving the resulting optimization problem. In each iteration, we solve a finite optimization problem and add one or more constraints. The proposed algorithm chooses a point at which the infinite constraints are violated to a degree rather than the violation being maximized. The organization of the rest of this paper is as follows. Section 2 provides some basic definitions to formulate the Brownian motion in the Vasicek-type forward interest rate model in terms of the deterministic perturbation. It shows that the Vasicek-type forward interest rate model can be solved via a nonlinear semi-infinite programming problem. Solution algorithms are developed in Section 3 for solving the resulting semi-infinite programming problem. The numerical results and comparison to some commonly used spline fitting methods are reported in Section 4. The paper is concluded in Section 5.

2. The Vasicek-type forward interest rate model with impulse perturbation

As mentioned in the previous section, in this paper a deterministic process is adopted to model the uncertainty in the interest rate behavior. It is assumed that the uncertainty is deterministic, which depends on the time to maturity τ. For convenience we denote the uncertainty as an integral function $w(\tau|t)$, and $\bar{w}(\tau|t),\ \underline{w}(\tau|t)$ are assumed to be the pre-assigned upper and lower bounds of $w(\tau|t)$, respectively, i.e.,

$$\underline{w}(\tau|t) \leqslant w(\tau|t) \leqslant \bar{w}(\tau|t). \tag{2}$$

In this case, the Vasicek-type forward interest rate model can be formulated as the following differential equation with uncertainty:

$$\frac{df(\tau|t)}{d\tau} = \alpha(t) + \beta(t)f(\tau|t) + w(\tau|t), \quad \forall \tau \geqslant t. \tag{3}$$

To specify the perturbation function $w(\tau|t)$, we introduce some notation and definitions.

Assume that for some present time t, there are M observed yields $\overline{R}_i$ with times to maturity $\tau_i,\ i = 1,2,\ldots,M$. Let $\widetilde{\aleph} \triangleq \{\tau_0, \tau_1, \ldots, \tau_M\}$, where $\tau_0 \triangleq t,\ \tau_{i-1} < \tau_i$, and $\aleph_i \triangleq [\tau_{i-1}, \tau_i],\ i = 1,2,\ldots,M$.

Definition 1 (*The observed Treasury yield*). The observed Treasury yield is defined as follows:

$$\overline{R}(\tau|t) \triangleq \overline{R}_i, \quad \forall \tau \in \aleph_i,\ i = 1,2,\ldots,M.$$

It is well known that using linear interpolation for all maturities between the observed yields may cause the forward rate yield curve being extremely bumpy and convex where it should be concave [36]. Therefore, in this work the observed Treasury yield is effectively defined as a step function.

Definition 2 (*The yield function*). The yield function is defined as the mean value of the integral of forward interest rates, i.e.,

$$R(\tau|t) \triangleq \frac{1}{\tau - t}\int_t^{\tau} f(s|t)\,ds. \tag{4}$$

Definition 3 (*The estimation error*). The estimation error is defined as the maximum absolute value of the difference of the yield function and the observed Treasury yield, i.e.,

$$\epsilon_i \triangleq \max_{\tau \in \aleph_i} |R(\tau|t) - \overline{R}(\tau|t)|, \quad i = 1,2,\ldots,M. \tag{5}$$

Definition 4 (*The impulse perturbation*). Let $w(\tau|t) \triangleq w_i(\tau|t),\ \forall \tau \in \aleph_i,\ i = 1,2,\ldots,M$. The impulse perturbation is defined to be

$$w_i(\tau|t) = k_i, \quad \forall \tau \in \aleph_i,\ i = 1,2,\ldots,M, \tag{6}$$

where $k_i \in \mathbb{R},\ i = 1,2,\ldots,M$, are constants and

$$\underline{w}_i(\tau|t) \leqslant w_i(\tau|t) \leqslant \bar{w}_i(\tau|t), \quad \forall \tau \in \aleph_i,\ i = 1,2,\ldots,M,$$

with $\underline{w}_i(\tau|t)$ and $\bar{w}_i(\tau|t),\ i = 1,2,\ldots,M$, are pre-assigned bounds for the perturbations.

The solution of the Vasicek-type forward interest rate model (3) with the impulse perturbation function defined in (6) can be derived from Theorem 1.

Theorem 1. *The instantaneous forward interest rate function of the Vasicek-type forward interest rate model (3) is given by*

$$f(\tau|t) = r(t)e^{\beta(t)(\tau-t)} + \frac{\alpha(t)}{\beta(t)}\left(e^{\beta(t)(\tau-t)} - 1\right) + \frac{1}{\beta(t)}\sum_{j=1}^{i-1}\left(e^{\beta(t)(\tau-\tau_{j-1})} - e^{\beta(t)(\tau-\tau_j)}\right)w_j(\tau|t) + \frac{\left(e^{\beta(t)(\tau-\tau_{i-1})} - 1\right)w_i(\tau|t)}{\beta(t)}, \quad \tau \in \aleph_i,\ i = 1,\ldots,M. \tag{7}$$

Proof. Multiply both sides with the integrating factor $e^{-\beta(t)(\tau-t)}$ for (3) we have

$$\left(\frac{df(\tau|t)}{d\tau}-\beta(t)f(\tau|t)\right)e^{-\beta(t)(\tau-t)}=\alpha(t)e^{-\beta(t)(\tau-t)}+w(\tau|t)e^{-\beta(t)(\tau-t)}.$$

Integrating each side from t to τ

$$\int_t^\tau (df(s|t)e^{-\beta(t)(s-t)})=\int_t^\tau (\alpha(t)e^{-\beta(t)(s-t)}+w(\tau|t)e^{-\beta(t)(s-t)})\,ds,$$

$$\begin{aligned}f(\tau|t)e^{-\beta(t)(\tau-t)}-f(t|t)e^{-\beta(t)(t-t)}&=\frac{\alpha(t)}{\beta(t)}(1-e^{-\beta(t)(\tau-t)})+\int_t^\tau w(s|t)e^{-\beta(t)(s-t)}\,ds=\frac{\alpha(t)}{\beta(t)}(1-e^{-\beta(t)(\tau-t)})+\int_t^{\tau_1} w_1(s|t)e^{-\beta(t)(s-t)}\,ds\\&\quad+\int_{\tau_1}^{\tau_2} w_2(s|t)e^{-\beta(t)(s-t)}\,ds+\cdots+\int_{\tau_{i-2}}^{\tau_{i-1}} w_{i-1}(s|t)e^{-\beta(t)(s-t)}\,ds+\int_{\tau_{i-1}}^{\tau} w_i(s|t)e^{-\beta(t)(s-t)}\,ds,\end{aligned}$$

$$\begin{aligned}f(\tau|t)e^{-\beta(t)(\tau-t)}&=r(t)+\frac{\alpha(t)}{\beta(t)}(1-e^{-\beta(t)(\tau-t)})+\frac{w_1(\tau|t)}{\beta(t)}(1-e^{-\beta(t)(\tau_1-t)})+\frac{w_2(\tau|t)}{\beta(t)}(e^{-\beta(t)(\tau_1-t)}-e^{-\beta(t)(\tau_2-t)})+\cdots\\&\quad+\frac{w_{i-1}(\tau|t)}{\beta(t)}(e^{-\beta(t)(\tau_{i-2}-t)}-e^{-\beta(t)(\tau_{i-1}-t)})+\frac{w_i(\tau|t)}{\beta(t)}(e^{-\beta(t)(\tau_{i-1}-t)}-1).\end{aligned}$$

Hence

$$f(\tau|t)=r(t)e^{\beta(t)(\tau-t)}+\frac{\alpha(t)}{\beta(t)}(e^{\beta(t)(\tau-t)}-1)+\frac{1}{\beta(t)}\sum_{j=1}^{i-1}(e^{\beta(t)(\tau-\tau_{j-1})}-e^{\beta(t)(\tau-\tau_j)})w_j(\tau|t)+\frac{(e^{\beta(t)(\tau-\tau_{i-1})}-1)w_i(\tau|t)}{\beta(t)},\quad \tau\in\aleph_i,\ i=1,\ldots,M.\qquad\square$$

It is well known that the yield function is one of the most important financial indicators in the theory of modern economics and finance. Substituting (7) into (4) yields the following result.

Theorem 2. *The yield function has the form*

$$\begin{aligned}R(\tau|t)&=\frac{1}{\tau-t}\left\{\frac{e^{\beta(t)(\tau-t)}-1}{\beta(t)}r(t)+\left(\frac{e^{\beta(t)(\tau-t)}-1}{\beta^2(t)}-\frac{\tau-t}{\beta(t)}\right)\alpha(t)+\sum_{j=1}^{i-1}\left(\frac{e^{\beta(t)(\tau-\tau_{j-1})}-e^{\beta(t)(\tau-\tau_j)}}{\beta^2(t)}-\frac{\tau_j-\tau_{j-1}}{\beta(t)}\right)w_j(\tau|t)\right.\\&\quad\left.+\left(\frac{e^{\beta(t)(\tau-\tau_{i-1})}-1}{\beta^2(t)}-\frac{\tau-\tau_{i-1}}{\beta(t)}\right)w_i(\tau|t)\right\},\quad \tau\in\aleph_i,\ i=1,2,\ldots,M.\end{aligned}\tag{8}$$

Proof

$$\begin{aligned}\frac{1}{\tau-t}\int_t^\tau f(s|t)\,ds&=\frac{1}{\tau-t}\int_t^\tau\left\{r(t)e^{\beta(t)(s-t)}+\frac{\alpha(t)}{\beta(t)}(e^{\beta(t)(s-t)}-1)+\sum_{j=1}^{i-1}(e^{\beta(t)(\tau-\tau_{j-1})}-e^{\beta(t)(\tau-\tau_j)})w_j(s|t)+\frac{(e^{\beta(t)(\tau-\tau_{i-1})}-1)}{\beta(t)}w_i(s|t)\right\}ds\\
&=\frac{1}{\tau-t}\left\{\frac{r(t)}{\beta(t)}e^{\beta(t)(s-t)}\Big|_t^\tau+\frac{\alpha(t)}{\beta^2(t)}(e^{\beta(t)(s-t)}-\beta(t)s)\Big|_t^\tau+\int_t^{\tau_1}\frac{e^{\beta(t)(s-t)}-1}{\beta(t)}w_1(s|t)\,ds+\int_{\tau_1}^{\tau_2}\left(\frac{e^{\beta(t)(s-t)}-e^{\beta(t)(s-\tau_1)}}{\beta(t)}w_1(s|t)\right.\right.\\
&\quad\left.+\frac{e^{\beta(t)(s-\tau_1)}-1}{\beta(t)}w_2(s|t)\right)ds+\cdots+\int_{\tau_{i-2}}^{\tau_{i-1}}\left(\sum_{j=1}^{i-2}\frac{e^{\beta(t)(s-\tau_{j-1})}-e^{\beta(t)(s-\tau_j)}}{\beta(t)}w_j(s|t)+\frac{e^{\beta(t)(s-\tau_{i-2})}-1}{\beta(t)}w_{i-1}(s|t)\right)ds\\
&\quad\left.+\int_{\tau_{i-1}}^{\tau}\left(\sum_{j=1}^{i-1}\frac{e^{\beta(t)(s-\tau_{j-1})}-e^{\beta(t)(s-\tau_j)}}{\beta(t)}w_j(s|t)+\frac{e^{\beta(t)(s-\tau_{i-1})}-1}{\beta(t)}w_i(s|t)\right)ds\right\}\\
&=\frac{1}{\tau-t}\left\{\frac{e^{\beta(t)(\tau-t)}-1}{\beta(t)}r(t)+\left(\frac{e^{\beta(t)(\tau-t)}-1}{\beta^2(t)}-\frac{\tau-t}{\beta(t)}\right)\alpha(t)+\left(\int_t^{\tau_1}\frac{e^{\beta(t)(s-t)}-1}{\beta(t)}w_1(s|t)\,ds\right.\right.\\
&\quad\left.+\int_{\tau_1}^{\tau_2}\frac{e^{\beta(t)(s-t)}-e^{\beta(t)(s-\tau_1)}}{\beta(t)}w_1(s|t)\,ds+\cdots+\int_{\tau_{i-2}}^{\tau_{i-1}}\frac{e^{\beta(t)(s-t)}-e^{\beta(t)(s-\tau_1)}}{\beta(t)}w_1(s|t)\,ds+\int_{\tau_{i-1}}^{\tau}\frac{e^{\beta(t)(s-t)}-e^{\beta(t)(s-\tau_1)}}{\beta(t)}w_1(s|t)\,ds\right)\\
&\quad+\left(\int_{\tau_1}^{\tau_2}\frac{e^{\beta(t)(s-\tau_1)}-1}{\beta(t)}w_2(s|t)\,ds+\cdots+\int_{\tau_{i-2}}^{\tau_{i-1}}\frac{e^{\beta(t)(s-\tau_1)}-e^{\beta(t)(s-\tau_2)}}{\beta(t)}w_2(s|t)\,ds\right.\\
&\quad\left.\left.+\int_{\tau_{i-1}}^{\tau}\frac{e^{\beta(t)(s-\tau_1)}-e^{\beta(t)(s-\tau_2)}}{\beta(t)}w_2(s|t)\,ds\right)+\cdots+\left(\int_{\tau_{i-1}}^{\tau}\frac{e^{\beta(t)(s-\tau_{i-1})}-1}{\beta(t)}w_i(s|t)\,ds\right)\right\}\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\tau - t}\left\{\frac{e^{\beta(t)(\tau-t)}-1}{\beta(t)} r(t) + (\frac{e^{\beta(t)(\tau-t)}-1}{\beta^2(t)} - \frac{\tau-t}{\beta(t)})\alpha(t) + \sum_{j=1}^{i-1}\left(\int_{\tau_{j-1}}^{\tau_j}\frac{e^{\beta(t)(s-\tau_{i-1})}-1}{\beta(t)} w_j(s|t)\,ds\right.\right.\\
&\quad \left.\left. + \int_{\tau_j}^{\tau}\frac{e^{\beta(t)(s-\tau_{j-1})}-e^{\beta(t)(s-\tau_j)}}{\beta(t)} w_j(s|t)\,ds\right) + \left(\int_{\tau_{i-1}}^{\tau}\frac{e^{\beta(t)(s-\tau_{i-1})}-1}{\beta(t)} w_i(s|t)\,ds\right)\right\}\\
&= \frac{1}{\tau - t}\left\{\frac{e^{\beta(t)(\tau-t)}-1}{\beta(t)} r(t) + \left(\frac{e^{\beta(t)(\tau-t)}-1}{\beta^2(t)} - \frac{\tau-t}{\beta(t)}\right)\alpha(t) + \sum_{j=1}^{i-1}\left(\frac{e^{\beta(t)(\tau-\tau_{j-1})}-e^{\beta(t)(\tau-\tau_j)}}{\beta^2(t)} - \frac{\tau_j-\tau_{j-1}}{\beta(t)}\right) w_j(\tau|t)\right.\\
&\quad \left. + \left(\frac{e^{\beta(t)(\tau-\tau_{i-1})}-1}{\beta^2(t)} - \frac{\tau-\tau_{i-1}}{\beta(t)}\right) w_i(\tau|t)\right\}. \qquad \square
\end{aligned}
$$

To shorten the mathematical formulas in (4), the following notations are introduced. Let

$$
\begin{aligned}
a_r(\tau,\beta(t)) &\triangleq \frac{1}{\tau-t}\left(\frac{e^{\beta(t)(\tau-t)}-1}{\beta(t)}\right),\\
a_\alpha(\tau,\beta(t)) &\triangleq \frac{1}{\tau-t}\left(\frac{e^{\beta(t)(\tau-t)}-1}{\beta^2(t)} - \frac{\tau-t}{\beta(t)}\right),\\
a_j(\tau,\beta(t)) &\triangleq \frac{1}{\tau-t}\left(\frac{e^{\beta(t)(\tau-\tau_{j-1})}-e^{\beta(t)(\tau-\tau_j)}}{\beta^2(t)} - \frac{\tau_j-\tau_{j-1}}{\beta(t)}\right),\\
\bar{a}_i(\tau,\beta(t)) &\triangleq \frac{1}{\tau-t}\left(\frac{e^{\beta(t)(\tau-\tau_{i-1})}-1}{\beta^2(t)} - \frac{\tau-\tau_{i-1}}{\beta(t)}\right).
\end{aligned} \tag{9}
$$

This work involves finding the impulse perturbation $w(\tau|t)$ that minimizes estimation errors. It leads to the following optimization problem.

Problem 1

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{M}\epsilon_i^2\\
\text{subject to} \quad & \overline{R}(\tau|t) \leqslant R(\tau|t) + \epsilon_i, \quad \forall\tau\in\aleph_i,\ i=1,2,\ldots,M,\\
& \overline{R}(\tau|t) \geqslant R(\tau|t) - \epsilon_i, \quad \forall\tau\in\aleph_i,\ i=1,2,\ldots,M,\\
& \underline{\alpha}(t) \leqslant \alpha(t) \leqslant \bar{\alpha}(t),\\
& \underline{\beta}(t) \leqslant \beta(t) \leqslant \bar{\beta}(t),\\
& \underline{r}(t) \leqslant r(t) \leqslant \bar{r}(t),\\
& \underline{w}_i(\tau|t) \leqslant w_i(\tau|t) \leqslant \bar{w}_i(\tau|t), \quad \forall\tau\in\aleph_i,\ i=1,2,\ldots,M,\\
& \epsilon_i \geqslant 0, \quad i=1,2,\ldots,M.
\end{aligned} \tag{10}
$$

Substituting (9) into Problem 1 leads to the following nonlinear programming problem.

Problem 2

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{M}\epsilon_i^2\\
\text{subject to} \quad & \overline{R}_i \leqslant a_r(\tau,\beta(t))r(t) + a_\alpha(\tau,\beta(t))\alpha(t) + \sum_{j=1}^{i-1} a_j(\tau,\beta(t))w_j(\tau|t) + \bar{a}_i(\tau,\beta(t))w_i(\tau|t) + \epsilon_i, \quad \forall\tau\in\aleph_i,\ i=1,2,\ldots,M,\\
& \overline{R}_i \geqslant a_r(\tau,\beta(t))r(t) + a_\alpha(\tau,\beta(t))\alpha(t) + \sum_{j=1}^{i-1} a_j(\tau,\beta(t))w_j(\tau|t) + \bar{a}_i(\tau,\beta(t))w_i(\tau|t) - \epsilon_i, \quad \forall\tau\in\aleph_i,\ i=1,2,\ldots,M,\\
& \underline{\alpha}(t) \leqslant \alpha(t) \leqslant \bar{\alpha}(t),\\
& \underline{\beta}(t) \leqslant \beta(t) \leqslant \bar{\beta}(t),\\
& \underline{r}(t) \leqslant r(t) \leqslant \bar{r}(t),\\
& \underline{w}_i(\tau|t) \leqslant w_i(\tau|t) \leqslant \bar{w}_i(\tau|t), \quad \forall\tau\in\aleph_i,\ i=1,2,\ldots,M,\\
& \epsilon_i \geqslant 0, \quad i=1,2,\ldots,M.
\end{aligned} \tag{11}
$$

Moreover, to be consistent with the theory of finance, some additional objectives are considered [19]:

Objective I. To ensure that the spot rate is instantaneously risk-free, i.e.,

$$\lim_{T\to t} R(T|t) = f(t|t) = r(t),$$

the spot rate $r(t)$ should be as close as possible to the observed Treasury yield $\overline{R}_1$ having the shortest term to maturity.

Objective II. To ensure that the spot forward rate tends to the long run mean, the mean reversion ratio, $-\frac{\alpha}{\beta}$, should be as close as possible to the yield $\overline{R}_M$ having the largest term to maturity.
Objective III. To provide a better fit to the empirical data, the perturbations acting on the forward interest rate curve should be as small as possible(stability).

It should be noted that in practice a moderately large simulation time is enough for characterizing the mean-reverting property of the Objective II (e.g. 53 weeks is considered in the numerical implementation of Ref. [19]) and the mean reversion is relatively fast, i.e. β is not too close to zero.

To specify the above objectives, the following notations are introduced:

$$\gamma_1 \triangleq (\overline{R}_1 - r(t))^2, \quad \gamma_2 \triangleq \left(\overline{R}_M + \frac{\alpha}{\beta}\right)^2, \quad \gamma_3 \triangleq \sum_{i=1}^{M} w_i^2(\tau|t), \tag{12}$$

where γ_1 is a measure of the distance of the yield of shortest observed maturity from the spot rate at the current time, γ_2 is the measure of the distance of the yield of the largest observed maturity to the mean reversion ratio, and γ_3 is the sum of the perturbation squares.

Applying (12) to Problem 2 generates the following problem.

Problem 3

$$\begin{aligned}
\min \quad & \gamma_1 + \gamma_2 + \gamma_3 + \theta \sum_{i=1}^{M} \epsilon_i^2 \\
\text{subject to} \quad & \overline{R}_i \leqslant a_r(\tau, \beta(t)) r(t) + a_\alpha(\tau, \beta(t))\alpha(t) + \sum_{j=1}^{i-1} a_j(\tau, \beta(t)) w_j(\tau|t) + \bar{a}_i(\tau, \beta(t)) w_i(\tau|t) + \epsilon_i, \quad \forall \tau \in \aleph_i,\ i = 1, 2, \ldots, M, \\
& \overline{R}_i \geqslant a_r(\tau, \beta(t)) r(t) + a_\alpha(\tau, \beta(t))\alpha(t) + \sum_{j=1}^{i-1} a_j(\tau, \beta(t)) w_j(\tau|t) + \bar{a}_i(\tau, \beta(t)) w_i(\tau|t) - \epsilon_i, \quad \forall \tau \in \aleph_i,\ i = 1, 2, \ldots, M, \\
& \underline{\alpha}(t) \leqslant \alpha(t) \leqslant \bar{\alpha}(t), \\
& \underline{\beta}(t) \leqslant \beta(t) \leqslant \bar{\beta}(t), \\
& \underline{r}(t) \leqslant r(t) \leqslant \bar{r}(t), \\
& \underline{w}_i(\tau|t) \leqslant w_i(\tau|t) \leqslant \bar{w}_i(\tau|t), \quad i = 1, 2, \ldots, M, \\
& \epsilon_i \geqslant 0, \quad i = 1, 2, \ldots, M,
\end{aligned}$$

where θ is a penalty coefficient of the model. It should be noted that Problem 3 is an ill-posed semi-infinite programming problem with finite variables, $\alpha,\ \beta,\ r,\ w_i,\ \epsilon_i,\ i = 1, 2, \ldots, M$, and infinitely many constraints. The numerical implementation for ill-posed semi-infinite programming problems has been recently discussed in [34].

3. An algorithm

In this work a cutting plane based algorithm is considered to effectively deal with the infinite number of constraints in Problem 3 [14,16,17]. Following the basic concept of the cutting plane approach, we can easily design an iterative algorithm which adds one or more constraints at a time for consideration until an optimal solution is identified. To be more specific, at the kth iteration, given subsets $N_i^k = \left\{\tau_1^i, \tau_2^i, \ldots, \tau_{p_i^k}^i\right\}$ and $\aleph_i^k = \left\{u_1^i, u_2^i, \ldots, u_{q_i^k}^i\right\}$ of $\aleph_i$, where $p_i^k,\ q_i^k \geqslant 1,\ i = 1, 2, \ldots, M$, we consider the following finite optimization problem:

Program SD^k

$$\begin{aligned}
\min \quad & \phi(\alpha, \beta, r, w, \epsilon) = \gamma_1 + \gamma_2 + \gamma_3 + \theta \sum_{i=1}^{M} \epsilon_i^2 \\
\text{subject to} \quad & \overline{R}_i \leqslant a_r(\tau_s^i, \beta(t)) r(t) + a_\alpha(\tau_s^i, \beta(t))\alpha(t) + \sum_{j=1}^{i-1} a_j(\tau_s^i, \beta(t)) w_j(\tau_s^i|t) + \bar{a}_i(\tau_s^i, \beta(t)) w_i(\tau_s^i|t) + \epsilon_i, \quad s = 1, 2, \ldots, p_i^k,\ i = 1, 2, \ldots, M, \\
& \overline{R}_i \geqslant a_r(u_l^i, \beta(t)) r(t) + a_\alpha(u_l^i, \beta(t))\alpha(t) + \sum_{j=1}^{i-1} a_j(u_l^i, \beta(t)) w_j(u_l^i|t) + \bar{a}_i(u_l^i, \beta(t)) w_i(u_l^i|t) - \epsilon_i, \quad l = 1, 2, \ldots, q_i^k,\ i = 1, \ldots, M, \\
& \underline{\alpha}(t) \leqslant \alpha(t) \leqslant \bar{\alpha}(t), \\
& \underline{\beta}(t) \leqslant \beta(t) \leqslant \bar{\beta}(t), \\
& \underline{r}(t) \leqslant r(t) \leqslant \bar{r}(t), \\
& \underline{w}_i(\tau|t) \leqslant w_i(\tau|t) \leqslant \bar{w}_i(\tau|t), \quad i = 1, 2, \ldots, M, \\
& \epsilon_i \geqslant 0, \quad i = 1, 2, \ldots, M,
\end{aligned}$$

where $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_M)$. Let F^k be the feasible region of Program SD^k. Suppose that $(\alpha^k, \beta^k, r^k, w^k, \epsilon^k)$ is an optimal solution of SD^k. We define the "constraint violation functions" as follows:

$$g_i^{k+1}(\tau) \triangleq \overline{R}_i - a_r(\tau, \beta^k(t)) r^k(t) - a_\alpha(\tau, \beta^k(t))\alpha^k(t) - \sum_{j=1}^{i-1} a_j(\tau, \beta^k(t)) w_j^k(\tau|t) - \bar{a}_i(\tau, \beta^k(t)) w_i^k(\tau|t) - \epsilon_i^k, \quad \tau \in \aleph_i,\ i = 1, 2, \ldots, M, \tag{13}$$

and

$$v_i^{k+1}(u) \triangleq a_r(u,\beta^k(t))r^k(t) + a_\alpha(u,\beta^k(t))\alpha^k(t) + \sum_{j=1}^{i-1} a_j(u,\beta^k(t))w_j^k(u|t) + \bar{a}_i(u,\beta^k(t))w_i^k(u|t) - \epsilon_i^k - \overline{R}_i, \quad u \in \aleph_i,\ i = 1,2,\ldots,M. \tag{14}$$

Since $\overline{R}_i,\ a_r,\ a_\alpha,\ a_j,\ \bar{a}_i$ are continuous over the compact set $\aleph_i,\ i = 1,2,\ldots,M$, the function $g_i^{k+1}(\tau)$ achieves its maximum over $\aleph_i,\ i = 1,2,\ldots,M$. A similar argument holds for the function $v_i^{k+1}(u),\ i = 1,2,\ldots,M$. Let $\tau^i_{p_i^k+1}$ and $u^i_{q_i^k+1}$ be such maximizers, $i = 1,2,\ldots,M$, and consider the values of $g_i^{k+1}(\tau^i_{p_i^k+1})$ and $v_i^{k+1}(u^i_{q_i^k+1}),\ i = 1,2,\ldots,M$. If the values are less than or equal to zero, then $(\alpha^k,\beta^k,r^k,w^k,\epsilon^k)$ becomes a feasible solution of Problem 3, and hence $(\alpha^k,\beta^k,r^k,w^k,\epsilon^k)$ is optimal for Problem 3 (because the feasible region F^k of Program SD^k is no smaller than the feasible region of Problem 3). Otherwise, we know that at least $\tau^i_{p_i^k+1} \notin N_i^k$ or $u^i_{q_i^k+1} \notin \aleph_i^k,\ i = 1,2,\ldots,M$. This background provides a foundation for us to outline a cutting plane algorithm for solving Problem 3.

CPSD algorithm

Initialization
Set $k = p_i^k = q_i^k = 1,\ i = 1,2,\ldots,M$; choose any $\tau_1^i,\ u_1^i \in \aleph_i,\ i = 1,2,\ldots,M$; set $N_i^1 = \{\tau_1^i\}$ and $\aleph_i^1 = \{u_1^i\},\ i = 1,2,\ldots,M$.

Step 1. Solve SD^k and obtain an optimal solution $(\alpha^k,\beta^k,r^k,w^k,\epsilon^k)$.
Step 2. Find a maximizer $\tau^i_{p_i^k+1}$ of $g_i^{k+1}(\tau)$ over $\aleph_i$ and a maximizer $u^i_{q_i^k+1}$ of $v_i^{k+1}(u)$ over $\aleph_i,\ i = 1,2,\ldots,M$.
Step 3. If $g_i^{k+1}\left(\tau^i_{p_i^k+1}\right) \leqslant 0$ and $v_i^{k+1}\left(u^i_{q_i^k+1}\right) \leqslant 0,\ i = 1,2,\ldots,M$, then stop with $(\alpha^k,\beta^k,r^k,w^k,\epsilon^k)$ being an optimal solution of Problem 3. Otherwise, go to Step 4.
Step 4. If $g_i^{k+1}\left(\tau^i_{p_i^k+1}\right) > 0$, then set $N_i^{k+1} \leftarrow N_i^k \bigcup \left\{\tau^i_{p_i^k+1}\right\},\ p_i^{k+1} \leftarrow p_i^k + 1$. Otherwise, set $N_i^{k+1} \leftarrow N_i^k,\ p_i^{k+1} \leftarrow p_i^k,\ i = 1,2,\ldots,M$.
Step 5. If $v_i^{k+1}\left(u^i_{q_i^k+1}\right) > 0$, then set $\aleph_i^{k+1} \leftarrow \aleph_i^k \bigcup \left\{u^i_{q_i^k+1}\right\},\ q_i^{k+1} \leftarrow q_i^k + 1$. Otherwise, set $\aleph_i^{k+1} \leftarrow \aleph_i^k,\ q_i^{k+1} \leftarrow q_i^k,\ i = 1,2,\ldots,M$.
Step 6. Set $k \leftarrow k + 1$ go to Step 1.

When Problem 3 has at least one feasible solution, it can be shown without much difficulty that the CPSD algorithm either terminates in a finite number of iterations with an optimal solution or generates a sequence of points $\{(\alpha^k,\beta^k,r^k,w^k,\epsilon^k),\ k = 1,2,\ldots\}$, which converges to an optimal solution $(\alpha^*,\beta^*,r^*,w^*,\epsilon^*)$, under some appropriate assumptions. However, for the above cutting plane algorithm, one major computation bottleneck lies in Step 2 of finding maximizers. Ideas for relaxing the requirement of finding global maximizers for different settings can be found in [15,13,35]. But the required computational work could still be a bottleneck. Here we propose a simple and yet very effective relaxation scheme which chooses points at which the infinite constraints are violated to a degree rather than the violation being maximized. The proposed algorithm is stated as follows.

Relaxed CPSD algorithm

Let $\delta > 0$ be a prescribed small number.
Initialization
Set $k = p_i^k = q_i^k = 1,\ i = 1,2,\ldots,M$; choose any $\tau_1^i,\ u_1^i \in \aleph_i,\ i = 1,2,\ldots,M$; set $N_i^1 = \{\tau_1^i\}$ and $\aleph_i^1 = \{u_1^i\},\ i = 1,2,\ldots,M$.

Step 1. Solve SD^k and obtain an optimal solution $(\alpha^k,\beta^k,r^k,w^k,\epsilon^k)$. Define $g_i^{k+1}(\tau)$ and $v_i^{k+1}(u),\ i = 1,2,\cdots,M$, according to (13) and (14), respectively.
Step 2. Find any $\tau^i_{p_i^k+1} \in \aleph_i$ such that $g_i^{k+1}\left(\tau^i_{p_i^k+1}\right) > \delta$, and $u^i_{q_i^k+1} \in \aleph_i$ such that $v_i^{k+1}\left(u^i_{q_i^k+1}\right) > \delta,\ i = 1,2,\ldots,M$.
Step 3. If such $\tau^i_{p_i^k+1}$ and $u^i_{q_i^k+1}$ do not exist, then output $(\alpha^k,\beta^k,r^k,w^k,\epsilon^k)$ as a solution. Otherwise, go to Step 4.
Step 4. If such $\tau^i_{p_i^k+1}$ exists, then set $N_i^{k+1} \leftarrow N_i^k \bigcup \left\{\tau^i_{p_i^k+1}\right\},\ p_i^{k+1} \leftarrow p_i^k + 1$. Otherwise, set $N_i^{k+1} \leftarrow N_i^k,\ p_i^{k+1} \leftarrow p_i^k,\ i = 1,2,\ldots,M$.
Step 5. If such $u^i_{q_i^k+1}$ exists, then set $\aleph_i^{k+1} \leftarrow \aleph_i^k \bigcup \left\{u^i_{q_i^k+1}\right\},\ q_i^{k+1} \leftarrow q_i^k + 1$. Otherwise, set $\aleph_i^{k+1} \leftarrow \aleph_i^k,\ q_i^{k+1} \leftarrow q_i^k,\ i = 1,2,\ldots,M$.
Step 6. Set $k \leftarrow k + 1$; go to Step 1.

Note that in Step 2, since no maximizer is required, the computational work can be greatly reduced. Also note that when δ is chosen to be sufficiently small, if the relaxed algorithm terminates in a finite number of iterations in Step 3, then an optimal solution is indeed obtained, assuming that the original Problem 3 is feasible. We now construct a convergence proof for the relaxed CPSD algorithm.

Theorem 3. *Given any $\delta > 0$, assume that there is a scalar $Q > 0$ such that $\|(\alpha,\beta,r,w,\epsilon)\| \leqslant \delta$ for each feasible solution $(\alpha,\beta,r,w,\epsilon)$ of SD^1 (Bounded Feasible Domain Assumption), then the relaxed CPSD algorithm terminates in a finite number of iterations.*

Proof. If the relaxed CPSD algorithm does not terminate in a finite number of iterations, then the algorithm generates an infinite sequence $\{(\alpha^k,\beta^k,r^k,w^k,\epsilon^k)\}_{k=1}^{\infty}$. We have

$$g_i^{k+1}\left(\tau^i_{p_i^k+1}\right) > \delta, \quad i = 1,2,\ldots,M,\ k = 1,2,\ldots, \tag{15}$$

and

$$v_i^{k+1}\left(u^i_{q_i^k+1}\right) > \delta, \quad i = 1,2,\ldots,M,\ k = 1,2,\ldots, \tag{16}$$

where $\tau^i_{p_i^k+1}$ and $u^i_{q_i^k+1}$ are generated by the relaxed CPSD algorithm.

Due to the Bounded Feasible Domain Assumption and the compactness of $\aleph_i$, $i = 1,2,\ldots,M$, there exists a subsequence $\{(\alpha^{k_j},\beta^{k_j},r^{k_j},w^{k_j},\epsilon^{k_j})\}$ of $\{(\alpha^k,\beta^k,r^k,w^k,\epsilon^k)\}$ such that $\lim_{j\to\infty}(\alpha^{k_j},\beta^{k_j},r^{k_j},w^{k_j},\epsilon^{k_j}) = (\alpha^*,\beta^*,r^*,w^*,\epsilon^*)$, $\lim_{j\to\infty}\tau^i_{p_i^{k_j}+1} = \tau^*$, and $\lim_{j\to\infty}u^i_{q_i^{k_j}+1} = u^*$. Consequently, by (15) and (16), we have

$$\overline{R}_i - a_r(\tau^*,\beta^*(t))r^*(t) - a_\alpha(\tau^*,\beta^*(t))\alpha^*(t) - \sum_{j=1}^{i-1} a_j(\tau^*,\beta^*(t))w_j^*(\tau^*|t) - \bar{a}_i(\tau^*,\beta^*(t))w_i^*(\tau^*|t) - \epsilon_i^* \geqslant \delta, \quad i = 1,2,\ldots,M,$$

and

$$a_r(u^*,\beta^*(t))r^*(t) + a_\alpha(u^*,\beta^*(t))\alpha^*(t) + \sum_{j=1}^{i-1} a_j(u^*,\beta^*(t))w_j^*(u^*|t) + \bar{a}_i(u^*,\beta^*(t))w_i^*(u^*|t) - \epsilon_i^* - \overline{R}_i \geqslant \delta, \quad i = 1,2,\ldots,M.$$

However, for each $\tau^i_{p_i^k}$ and $u^i_{q_i^k}$, $i = 1,2,\ldots,M$, $k = 1,2,\ldots,$

$$\overline{R}_i - a_r\left(\tau^i_{p_i^k},\beta^l(t)\right)r^l(t) - a_\alpha\left(\tau^i_{p_i^k},\beta^l(t)\right)\alpha^l(t) - \sum_{j=1}^{i-1} a_j\left(\tau^i_{p_i^k},\beta^l(t)\right)w_j^l\left(\tau^i_{p_i^k}|t\right) - \bar{a}_i\left(\tau^i_{p_i^k},\beta^l(t)\right)w_i^l\left(\tau^i_{p_i^k}|t\right) - \epsilon_i^l \leqslant 0, \quad i = 1,2,\ldots,M,\ \forall l \geqslant k,$$

and

$$a_r\left(u^i_{q_i^k},\beta^l(t)\right)r^l(t) + a_\alpha\left(u^i_{q_i^k},\beta^l(t)\right)\alpha^l(t) + \sum_{j=1}^{i-1} a_j\left(u^i_{q_i^k},\beta^l(t)\right)w_j^l\left(u^i_{q_i^k}|t\right) + \bar{a}_i\left(u^i_{q_i^k},\beta^l(t)\right)w_i^l\left(u^i_{q_i^k}|t\right) - \epsilon_i^l - \overline{R}_i \leqslant 0, \quad i = 1,2,\ldots,M,\ \forall l \geqslant k.$$

Therefore, for any fixed k, as the subsequence $\{(\alpha^{k_j},\beta^{k_j},r^{k_j},w^{k_j},\epsilon^{k_j})\} \to (\alpha^*,\beta^*,r^*,w^*,\epsilon^*)$, we see that

$$\overline{R}_i - a_r\left(\tau^i_{p_i^k},\beta^*(t)\right)r^*(t) - a_\alpha\left(\tau^i_{p_i^k},\beta^*(t)\right)\alpha^*(t) - \sum_{j=1}^{i-1} a_j\left(\tau^i_{p_i^k},\beta^*(t)\right)w_j^*\left(\tau^i_{p_i^k}|t\right) - \bar{a}_i\left(\tau^i_{p_i^k},\beta^*(t)\right)w_i^*\left(\tau^i_{p_i^k}|t\right) - \epsilon_i^* \leqslant 0, \quad i = 1,2,\ldots,M,$$

and

$$a_r\left(u^i_{q_i^k},\beta^*(t)\right)r^*(t) + a_\alpha\left(u^i_{q_i^k},\beta^*(t)\right)\alpha^*(t) + \sum_{j=1}^{i-1} a_j\left(u^i_{q_i^k},\beta^*(t)\right)w_j^*\left(u^i_{q_i^k}|t\right) + \bar{a}_i\left(u^i_{q_i^k},\beta^*(t)\right)w_i^*\left(u^i_{q_i^k}|t\right) - \epsilon_i^* - \overline{R}_i \leqslant 0, \quad i = 1,2,\ldots,M.$$

Since the above expression is true for all k, we have

$$\overline{R}_i - a_r(\tau^*,\beta^*(t))r^*(t) - a_\alpha(\tau^*,\beta^*(t))\alpha^*(t) - \sum_{j=1}^{i-1} a_j(\tau^*,\beta^*(t))w_j^*(\tau^*|t) - \bar{a}_i(\tau^*,\beta^*(t))w_i^*(\tau^*|t) - \epsilon_i^* \leqslant 0, \quad i = 1,2,\ldots,M,$$

and

$$a_r(u^*,\beta^*(t))r^*(t) + a_\alpha(u^*,\beta^*(t))\alpha^*(t) + \sum_{j=1}^{i-1} a_j(u^*,\beta^*(t))w_j^*(u^*|t) + \bar{a}_i(u^*,\beta^*(t))w_i^*(u^*|t) - \epsilon_i^* - \overline{R}_i \leqslant 0, \quad i = 1,2,\ldots,M,$$

which contradicts the facts that

$$\overline{R}_i - a_r(\tau^*,\beta^*(t))r^*(t) - a_\alpha(\tau^*,\beta^*(t))\alpha^*(t) - \sum_{j=1}^{i-1} a_j(\tau^*,\beta^*(t))w_j^*(\tau^*|t) - \bar{a}_i(\tau^*,\beta^*(t))w_i^*(\tau^*|t) - \epsilon_i^* \geqslant \delta, \quad i = 1,2,\ldots,M,$$

and

$$a_r(u^*,\beta^*(t))r^*(t) + a_\alpha(u^*,\beta^*(t))\alpha^*(t) + \sum_{j=1}^{i-1} a_j(u^*,\beta^*(t))w_j^*(u^*|t) + \bar{a}_i(u^*,\beta^*(t))w_i^*(u^*|t) - \epsilon_i^* - \overline{R}_i \geqslant \delta, \quad i = 1,2,\ldots,M.$$

The theorem is proved. □

4. Numerical results

In this section, the features of the proposed method are tested using a set of real data and compared with the smoothing spline method [2], the cubic smoothing spline method [27], and the maximum smoothing spline method [20]. The numerical experiments are performed on the Intel Pentium 4 3.0 GHz under the Windows XP Professional SP2 operating system. The data USFR011C Currency from 2005-10-18 to 2006-04-28 (81 Weekly) is employed for analysis. The data providers are Bloomberg Financial Markets. The US represent US dollar, FR means the forward interest rate and 011 is the data number. This data is a single Treasury yield curve observed on 2005-10-18 with maturities ranging from 0 to 81 weeks. Since the long data length chosen would be time-consuming in the implementation, only 81 observation periods are considered in our numerical experiments. Moreover, in our implementation, the parameter β is pre-assigned for simplifying the

Table 1
The initial values and bounds of the parameters of the Vasicek-type forward interest rate model.

	Initial value	Lower bound	Upper bound
$r(t)$	0	0	∞
α	0	0	∞
β	Shown in Table 2	$-\infty$	0
$w(\tau\|t)$	0	$-\infty$	∞

Table 2
The numerical results for different β, assuming $\theta = 10^5$.

β	ϵ_{max}	ϵ_{mean}	$r(t)$	$\alpha(t)$	R_{max}
−0.01	0.002502	0.000650	0.000110	0.0083	0.045393
−0.1	0.002018	0.000634	0.000676	0.1275	0.045522
−0.5	0.002019	0.000628	0.002842	0.1757	0.045524
−1	0.002019	0.000628	0.005869	0.1978	0.045482
−2	0.002021	0.000629	0.012781	0.2396	0.045395
−25	0.002342	0.000705	0.036659	1.1712	0.043806
−50	0.003087	0.000856	0.03765	2.1082	0.042938
−100	0.004038	0.001123	0.0376	4.1583	0.042253

Table 3
The numerical results for different θ, assuming $\beta = -0.5$.

θ	ϵ_{max}	ϵ_{mean}	$r(t)$	$\alpha(t)$	R_{max}
1	0.018076	0.0037728	0.0078197	0.05245	0.05767
10^2	0.011395	0.0017305	0.0066032	0.06214	0.05099
10^3	0.008794	0.0014077	0.0060641	0.06924	0.04839
10^4	0.005429	0.0011596	0.0049584	0.09912	0.04503
10^6	0.002019	0.0006285	0.0028426	0.17573	0.04552
10^7	0.001777	0.0004588	0.0023745	0.19137	0.04596
10^8	0.001596	0.0003749	0.0025354	0.17296	0.04610

Table 4
The numerical results for Objectives I–III with different β, assuming $\theta = 10^5$.

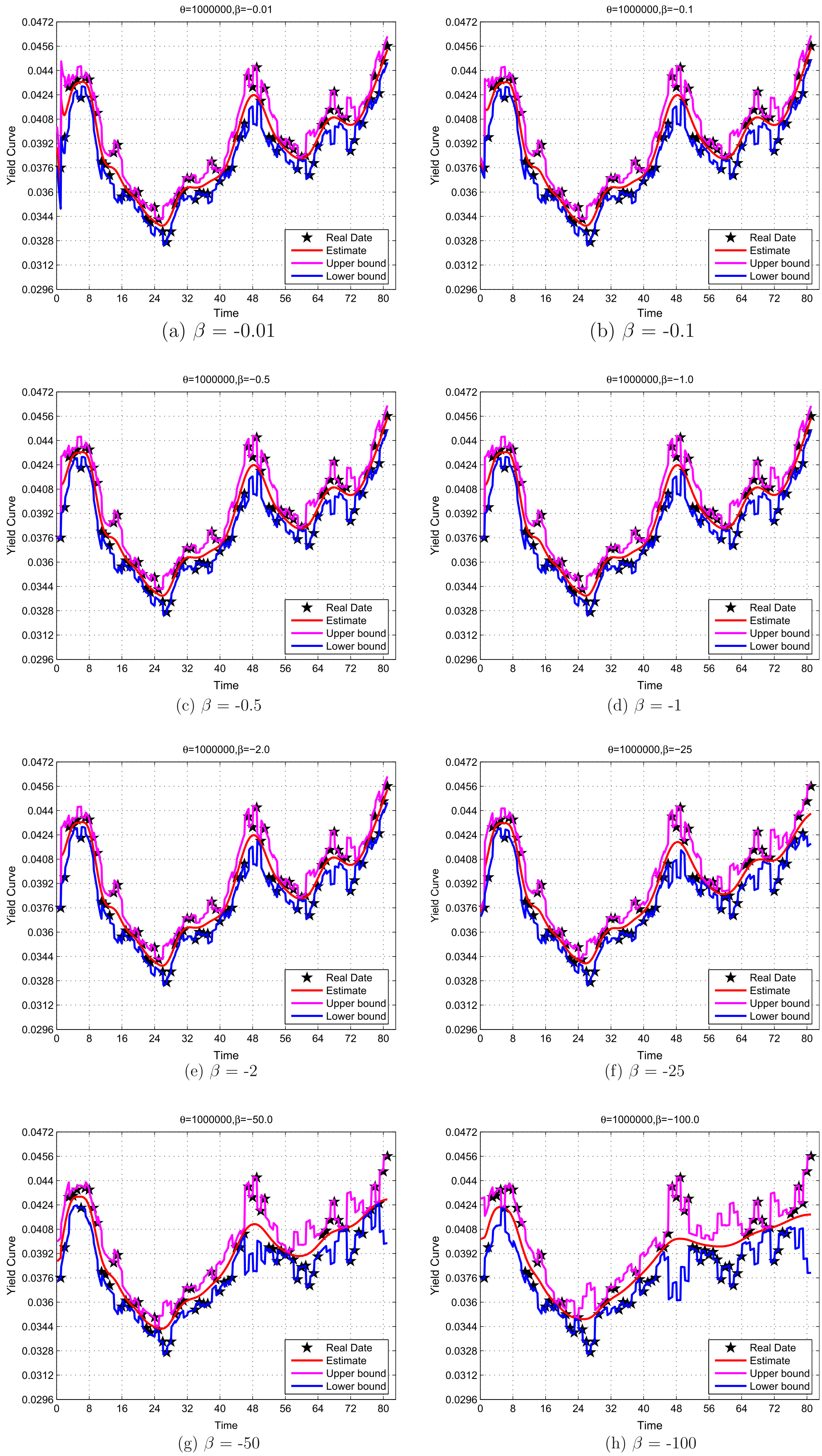
β	γ_1	γ_2	γ_3	$\frac{-\alpha}{\beta}$
−0.01	0.000856031	0.00119716	25.732494	0.011000
−0.1	0.001363382	1.51361348	19.724271	1.275890
−0.5	0.001208078	0.09355596	19.827829	0.351469
−1	0.001006818	0.02319035	19.946171	0.197883
−2	0.000616005	0.00551062	20.195738	0.119833
−25	0.000000886	0.00000155	35.794091	0.046847
−50	0.000000003	0.00001180	55.821307	0.042164
−100	0.000000000	0.00001613	82.162230	0.041582

Table 5
The numerical results for Objectives I–III with different θ, assuming $\beta = -0.5$.

θ	γ_1	γ_2	γ_3	$\frac{-\alpha}{\beta}$
1	0.00088804	0.00352836	0.00029404	0.105
10^2	0.00096080	0.00619183	0.00905942	0.12428818
10^3	0.00099451	0.00862762	0.07173936	0.13848499
10^4	0.00106547	0.02330265	0.65305515	0.19825206
10^6	0.00120807	0.09355597	19.82782998	0.35146920
10^7	0.00124083	0.11365750	79.90312096	0.38273128
10^8	0.00122952	0.09018912	387.11581104	0.34591505

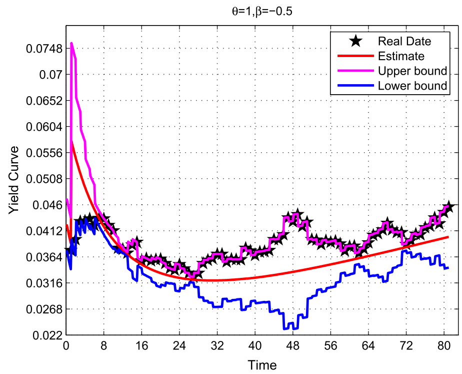
Table 6
The yield curve variance for different methods.

	The yield curve variance
Smoothing spline	8.86e−006
Cubic smoothing spline	8.89e−006
Maximum smoothing spline method	9.14e−006
Deterministic approach	4.97e−006

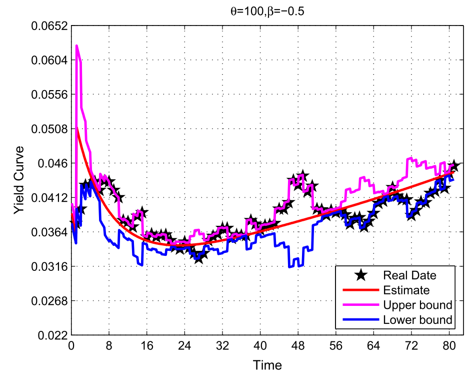


(a) β = -0.01
(b) β = -0.1
(c) β = -0.5
(d) β = -1
(e) β = -2
(f) β = -25
(g) β = -50
(h) β = -100

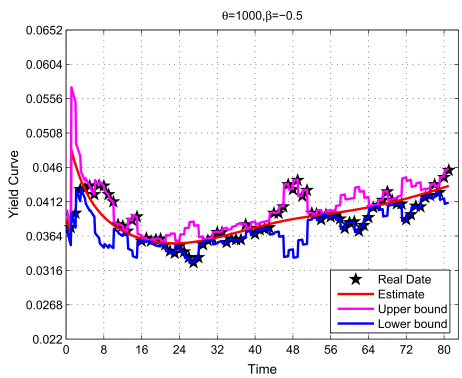
Fig. 1. Yield curves for different values of β.



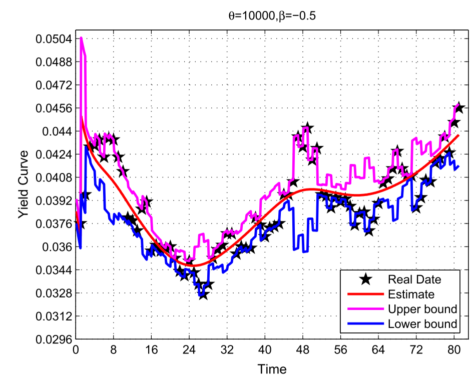
(a) $\theta = 1$



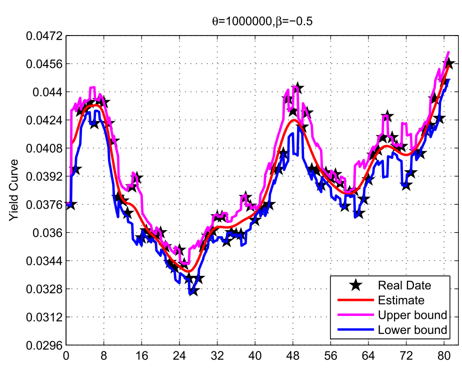
(b) $\theta = 100$



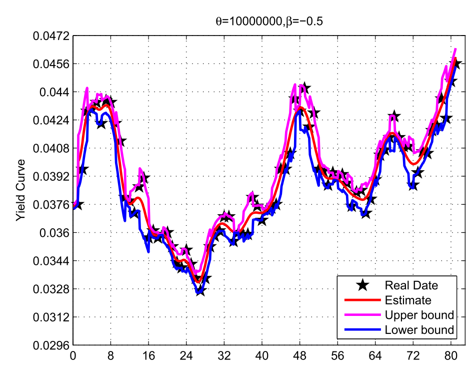
(c) $\theta = 1000$



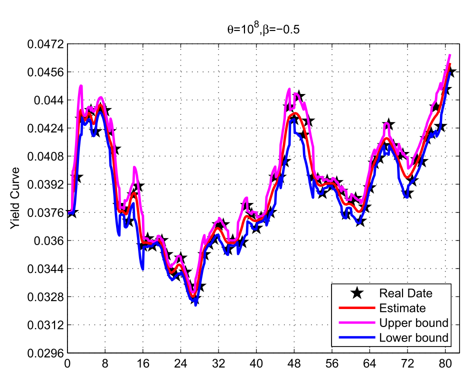
(d) $\theta = 10000$



(e) $\theta = 1000000$

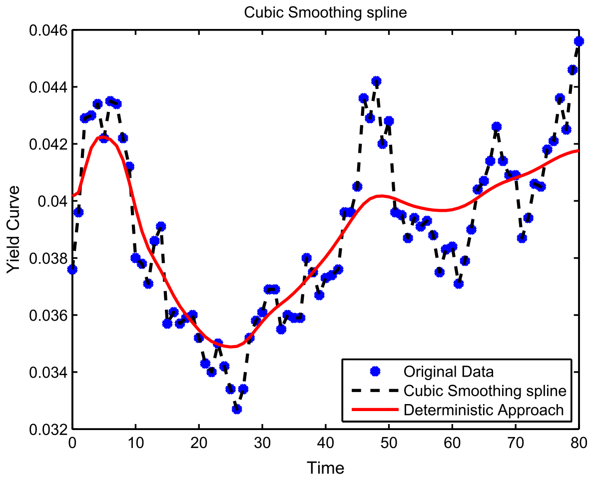


(f) θ =10000000

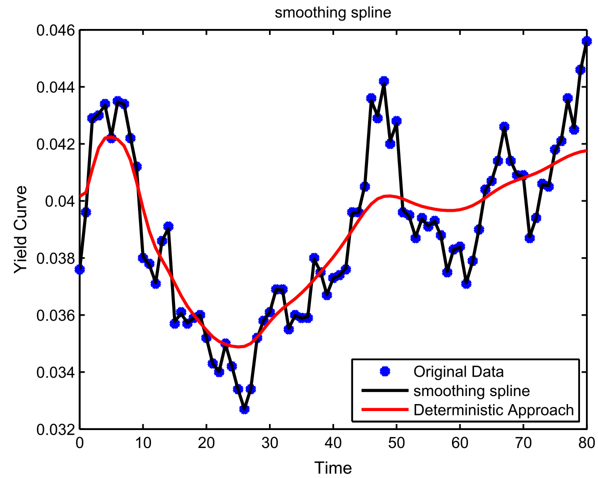


(g) θ =100000000

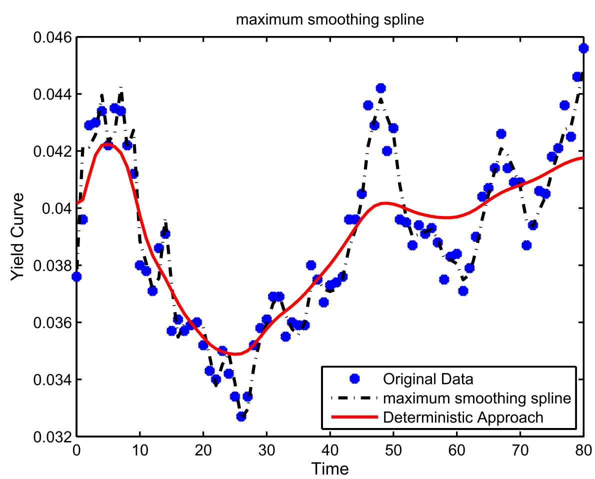
Fig. 2. Yield curves for different values of θ.



(a) Cubic smoothing spline method



(b) Smoothing spline method



(c) Maximum smoothing spline method

Fig. 3. The comparison of the yield curves for different spline methods and the deterministic approach, assuming $\theta = 10^5$ and $\beta = -100$.

constraints in Problem 3. The initial values and bounds of the parameters of the Vasicek-type forward interest rate model are listed in Table 1. The numerical results for different β with fixed $\theta = 10^5$ are shown in Tables 2 and 4. Tables 3 and 5 show the numerical results for different θ with fixed $\beta = -0.5$. In Tables 2 and 3, R_{max} denotes the maximum value of forward rate yield function, ϵ_{max} denotes the maximum error of the estimated value, i.e., $\epsilon_{max} \triangleq \max_{i=1,2,\ldots,M} \epsilon_i$, and ϵ_{mean} denotes the mean error of the estimate value, i.e., $\epsilon_{mean} \triangleq \frac{1}{M}\sum_{i=1}^{M} \epsilon_i$. Table 6 compares the yield curve variance for the smoothing spline method, the cubic smoothing spline method, the maximum smoothing spline method, and our approach. The result illustrates that our approach generates the yield function with smaller oscillation. Moreover, Fig. 1a–h show the estimates of yield curves for different values of β with fixed $\theta = 10^5$. Fig. 2a–g shows the estimates of yield curves for different values of θ with fixed $\beta = -0.5$. Fig. 3a–c shows the compares of the yield curves for different spline methods and the approach used in this paper, assuming $\beta = -100$ and $\theta = 10^5$. It should be noted that when β is chosen to be large enough, the value of the mean reversion value will tend to a certain value. Moreover, when θ is chosen to be large enough, the resulting forward interest rate and yield functions are highly smooth.

5. Conclusions

The Vasicek-type forward interest rate model with impulse perturbation has been studied. The concept of deterministic perturbation is adopted to deal with the random behavior of interest rate variation. It shows that the solution of the Vasicek-type forward interest rate model can be obtained by solving a nonlinear semi-infinite programming problem. A relaxed cutting plane algorithm is then proposed for solving the resulting optimization problem. In each iteration, we solve a finite optimization problem and add one or more constraints. The proposed algorithm chooses a point at which the infinite constraints are violated to a degree rather than the violation being maximized. Compared to some commonly used spline fitting methods, our approach generates the yield functions with minimal fitting errors and small oscillations.

Acknowledgments

The author would like to thank professor Fang-Bo Yeh and Huang-Nan Huang for their very constructive and valuable suggestions.

References

[1] Y. Ait-Sahalia, Transition densities for interest rate and other nonlinear diffusions, Journal of Finance 54 (1999) 1361–1395.
[2] C. de Boor, A Practical Guide to Splines, Springer-Verlag, New York, 1978.
[3] S.J. Brown, P.H. Dybvig, The empirical implications of the Cox, Ingersoll, Ross theory of the term structure of interest rates, Journal of Finance 41 (1986) 617–630.
[4] R.H. Brown, S.M. Schaefer, The term structure of real interest rates and the Cox, Ingersoll, and Ross model, Journal of Financial Economics 35 (1994) 3–42.
[5] W.T. Carleton, I.A. Cooper, Estimation and uses of the term structure of interest rates, Journal of Finance 31 (1976) 1067–1083.
[6] D.R. Chambers, W.T. Carleton, D. Waldman, A new approach to estimation of the term structure of interest rates, Journal of Financial and Quantitative Analysis 19 (1984) 233–252.
[7] K.C. Chan, G.A. Karolyi, F.A. Longstaff, A.B. Sanders, An empirical comparison of alternative models of the short-term interest rate, Journal of Finance 47 (1992) 1209–1227.
[8] G. Courtadon, The pricing of options on default-free bonds, Journal of Financial and Quantitative Analysis 17 (1982) 75–100.
[9] J.C. Cox, J.E. Ingersoll, S.A. Ross, A theory of the term structure of interest rates, Econometrica 53 (1985) 385–407.
[10] L.U. Dothan, On the term structure of interest rates, Journal of Financial Economics 6 (1978) 59–69.
[11] D. Duffie, K.J. Singleton, Simulated moments estimation of Markov models of asset prices, Econometrica 61 (1993) 929–952.
[12] D. Florens-Zmirou, On estimating the diffusion coefficient from discrete observations, Journal of Applied Probability 30 (1993) 790–804.
[13] P.R. Gribik, A central cutting plane algorithm for SIP, in: R. Hettich (Ed.), Semi-infinite Programming, Springer-Verlag, New York, 1979, pp. 66–82.
[14] S.A. Gustafson, K.O. Kortanek, Numerical solution of a class of semi-infinite programming problems, Naval Research Logistics Quarterly 20 (2) (1973) 477–504.
[15] S.A. Gustafson, K.O. Kortanek, Numerical solution of a class of convex programs, in: R. Henn, H.P. Kunzi, H. Schubert (Eds.), Methods of Operations Research, Verlag Anton Hain, Germany, 1973, pp. 138–149.
[16] S.A. Gustafson, K.O. Kortanek, Semi-infinite programming and applications, in: A. Bachem, M. Grotschel, B. Korte (Eds.), Mathematical Programming: State of the Art, Springer-Verlag, New York, 1983, pp. 132–157.
[17] R. Hettich, K.O. Kortanek, Semi-infinite programming: Theory, methods, and applications, SIAM Review 35 (3) (1993) 380–429.
[18] I. Karatzas, S. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, New York, 1991.
[19] K.O. Kortanek, V.G. Medvedev, Building and Using Dynamic Interest Rate Models, John Wiley and Sons, New Jersey, 2001.
[20] K.G. Lim, Q. Xiao, Computing maximum smoothness forward rate curves, Statistics and Computing 12 (2002) 275–279.
[21] B.H. Lin, Fitting the term structure of interest rates for Taiwanese government bonds, Journal of Multinational Financial Management 9 (1999) 331–352.
[22] J.H. McCulloch, A theory of the term structure of interest rates, Journal of Business 44 (1971) 19–31.
[23] J.F.J. de Munnik, P.C. Schotman, Cross-sectional versus time series estimation of the term structure models: Empirical results for the Dutch bond market, Journal of Banking and Finance 18 (1994) 997–1025.
[24] C. Nelson, A. Siegel, Parsimonious modeling of yield curves, Journal of Business 60 (1987) 473–489.
[25] K. Nowman, Gaussian estimation of single-factor continuous time models of the term structure of interest rates, Journal of Finance 52 (1997) 1695–1703.
[26] T.M. Pham, Estimation of the term structure of interest rates: An international perspective, Journal of Multinational Financial Management 8 (1998) 265–283.
[27] C.H. Reinsch, Smoothing by spline functions, Numerische Mathematik 10 (1967) 177–183.
[28] S.M. Schaefer, Measuring a tax-specific term structure of interest rates in the market for British government securities, The Economic Journal 91 (1981) 415–438.
[29] P. Sercu, X. Wu, P.C. Schotman, The information content in bond model residuals: An empirical study on the Belgian bond market, Journal of Banking and Finance 21 (1997) 685–720.
[30] M.S. Staffa, Un approccio dinamico per il moto dello Spot rate in base and osservazoni sulla Curva dei Rendimenti: Il caso del mercato italiano, Ph.D. Thesis, Universiatá degli Studi di Roma – La Sapienza, Dipartimento di Scienze Attuariali e Finanziarie, 1999.
[31] R.C. Stapleton, M.G. Subrahmanyam, Analysis and valuation of interest rate options, Journal of Banking and Finance 17 (1993) 1079–1095.
[32] J.M. Steeley, Estimating the gilt-edged term structure: Basis splines and confidence intervals, Journal of Business Finance and Accounting 18 (1991) 513–529.
[33] G.C. Stephen, The case of the negative nominal interest rates: New estimates of the term structure of interest rates during the Great depression, Journal of Political Economy 96 (1988) 1111–1141.
[34] R. Tichatschke, A. Kaplan, T. Voetmann, M. Böhm, Numerical treatment of an asset price model with non-stochastic uncertainty, TOP Sociedad de Estadistíca e Investigación Operativa 10 (2002) 1–50.
[35] R. Tichatschke, V. Nebeling, A cutting plane algorithm for solving quadratic semi-infinite programs, Optimization 19 (1988) 803–817.
[36] B. Tuckman, Fixed Income Securities: Tools for Today's Markets, Wiley, New York, 2002.
[37] O.A. Vasicek, An equilibrium characterization of the term structure, Journal of Financial Economics 5 (1977) 177–188.
[38] O.A. Vasicek, H.G. Fong, Term structure modeling using exponential splines, Journal of Finance 37 (1982) 339–356.