Comparative Analysis of Zero Coupon Yield Curve Estimation Methods Using JGB Price Data

Kentaro Kikuchi and Kohei Shintani

This paper conducts a comparative analysis of the diverse methods for estimating the Japanese government bond (JGB) zero coupon yield curve (hereafter, zero curve) according to the criteria that estimation methods should meet. Previous studies propose many methods for estimating the zero curve from the market prices of coupon-bearing bonds. In estimating the JGB zero curve, however, an undesirable method may fail to accurately grasp the features of the zero curve. To select an appropriate estimation method for the JGB, we set the following criteria for the zero curve: (1) estimates should not fall below zero; (2) estimates should not take abnormal values; (3) estimates should have a good fit to market prices; and (4) the zero curve should have little unevenness. The method that meets these criteria enables us to estimate the zero curve with a good fit to JGB market prices and a proper interpolation to grasp the features of the zero curve. Based on our analysis, we conclude that the method proposed in Steeley (1991) is the most appropriate in light of the criteria for the JGB price data. In fact, the zero curve based on this method can fully capture the characteristics of the JGB zero curve in a prolonged period of accommodative monetary policy.

Keywords: Coupon-bearing government bond; Zero coupon yield; Piecewise polynomial function
JEL Classification: C13, C14, G12

Kentaro Kikuchi: Deputy Director and Economist, Institute for Monetary and Economic Studies, Bank of Japan (E-mail: kentarou.kikuchi@boj.or.jp)
Kohei Shintani: Economist, Institute for Monetary and Economic Studies, Bank of Japan. Currently in the Personnel and Corporate Affairs Department (studying at the University of Illinois at Urbana-Champaign)

The authors would like to thank Yukio Muromachi (Tokyo Metropolitan University), the participants in the JAFEE 35th Conference, the participants in the Tokyo Metropolitan University Finance Seminar, and staff of the Bank of Japan (BOJ) for their useful comments. We would also like to thank Nikkei Media Marketing, Inc. for granting consent for the publication of zero curve data estimated using JGB price data sourced from that company's NEEDS service. All the remaining errors are our own. Views expressed in this paper are those of the authors and do not necessarily reflect the official views of the BOJ.

I. Introduction

The zero coupon yield (hereafter, zero yield) is defined as the yield to maturity of discount bonds. This is used to calculate the present value of future cash flow at any point of time. The zero coupon yield curve (hereafter, zero curve) that connects zero yields with different maturities enables us to calculate the present value of any financial product. Furthermore, using this we can conduct a comparative analysis between interest rates with different maturities. Therefore, the zero curve is essential for researchers, analysts, policymakers, and other users.

If bonds with all maturities are traded, we can calculate the zero curve from their market prices. However, this is usually not the case in the actual market. Therefore, we must estimate zero yields for maturities of bonds traded in the market and interpolate them to obtain zero yields for maturities of bonds that are not traded.[1] If discount bonds are traded, then zero yields with their maturities can be directly derived from their market prices, and thus the remaining problem is how zero yields for maturities of non-traded bonds are interpolated. However, in the Japanese government bond (JGB) market, there are few discount bonds with remaining maturities of one year or more. Thus, it is difficult to derive the JGB zero curve from the market prices of discount bonds,[2] and the zero curve must be estimated from the market prices of coupon-bearing government bonds, which have a greater variety of maturities.

Diverse methods for estimating the zero curve from the market prices of coupon-bearing government bonds have been proposed in previous studies. Representative methods include (1) the piecewise polynomial method (McCulloch [1971, 1975], Steeley [1991], etc.), which models the discount function with piecewise polynomials; (2) the non-parametric method (Tanggaard [1997], etc.), which does not assume any specific structure for the discount function; (3) the polynomial method (Schaefer [1981], etc.), which models the discount function with polynomials; and (4) the parsimonious function method (Nelson and Siegel [1987], Svensson [1995], etc.), which assumes specific functional forms for the zero yield or instantaneous forward rate. In utilizing the zero curve for an analysis on interest rates, the zero curve is estimated by one of these four estimation methods. Bank for International Settlements (2005) summarizes the methods used by the central banks in estimating their government bond zero curves, and each central bank uses one of the above methods.

1. In Japan, discount bonds presently issued by the government have a maturity of one year or less. However, the Act on Book-Entry Transfer of Company Bonds, Shares, etc., permits separation of the principal and interest portions of all fixed-interest government bonds (excluding inflation-indexed bonds and government bonds for individual investors) issued after January 27, 2003 (that is, it makes these bonds “strippable”). This makes it possible to trade discount government bonds with a maturity of one year or more on the secondary market. However, we must note that the liquidity of the market is poor compared with the fixed-interest government bond market.
2. JGB interest rate data can be obtained from information vendors and the Ministry of Finance homepage. However, as far as the authors know, the methodology for the zero curve estimates provided by the information vendors has not been disclosed. Moreover, the interest rates released by the Ministry of Finance are the yields to maturity of fixed-interest government bonds on a semiannual compounded basis, not the zero yield. In addition to these sources, some researchers release data used for their research on their homepages. For example, Johns Hopkins University Professor Jonathan Wright releases zero coupon curve estimates for each country including Japan on a monthly basis on his homepage. He notes that these estimates are based on the Svensson (1995) method handled in this paper, but the data sources are not identified.

When we estimate the JGB zero curve, selecting a method without careful consideration might result in the estimation of a curve that does not grasp the characteristics of the JGB yield curve. Moreover, research and analysis using such a zero curve could lead to the wrong conclusions. To avoid such problems and select an estimation method that can accurately grasp the characteristics of the JGB yield curve, this paper compares several estimation methods proposed in previous studies.

Prior papers that compare multiple zero curve estimation methods include Ioannides (2003) for U.K. government bonds and Kalev (2004) for Australian government bonds. However, good estimation methods in previous studies are not necessarily good for the JGB market, since developments in government bond market prices and market practices differ from country to country. Previous studies and surveys on the estimation of the JGB zero curve include Komine *et al*. (1989), Oda (1996), Inui and Muromachi (2000), and Kawasaki and Ando (2005). However, only Komine *et al*. (1989) compare multiple estimation methods. Since their paper uses JGB price data from the second half of the 1980s to compare five estimation methods, we must note that there is a great difference between the JGB market environments in the 1980s and since the late 1990s.[3] Accordingly, using JGB price data from 1999 to 2010, we compare the representative estimation methods proposed in previous studies and select an estimation method that can grasp the characteristics of the JGB yield curve.

For JGB yield curves since 1999, yield curves under the zero interest rate policy and the quantitative easing policy are distinctive. As a feature of yield curves during these periods, we can point out that the yield curve has a flat shape near zero at the short-term maturities. Some estimation methods cannot grasp this kind of curve shape, and sometimes estimate zero yields below zero. Thus, it is necessary to compare the estimation methods for the JGB zero curve. In this paper, we set the criteria to make an appropriate selection of the estimation method; we then select the optimal estimation method based on the criteria. This approach has not been tried in previous studies. Specifically, we first reject inappropriate methods based on the following criteria: (1) zero yield estimates should not fall below zero; and (2) zero yield estimates should not take abnormal values. Next, from among the remaining estimation methods, the most desirable method is chosen based on the criteria of (3) a good fit to JGB market prices and (4) little unevenness in the zero curve. As a result of this selection process, the estimation method in Steeley (1991) is selected.

The remainder of this paper is organized as follows. Section II explains the zero curve estimation methods. Section III presents the selection criteria for the zero curve estimation methods and uses these criteria to compare several representative estimation methods proposed in the prior studies. Section IV clarifies the characteristics of the method in Steeley (1991) when applied to the JGB market through comparisons with other estimation methods. Section V presents our conclusion.

As reference materials to help readers reproduce the estimation methods, Appendix 1 summarizes the JGB market conventions required to calculate theoretical JGB prices

3. In addition to differences in the interest rate term structure itself including the level and the curve shape, as a market practice, there was the so-called benchmark issue in the 1980s. There was intensive trading of the benchmark issue, while there were few trades of other issues. The concentration of trading on the benchmark issue declined from the mid-1990s, and the benchmark issue disappeared in the late 1990s.

such as the definition of the timing of when cash flows are paid, the method of calculating the number of days until cash flows are paid, and the method of calculating accrued interest. Appendix 2 explains the details of the estimation algorithm in Steeley (1991). For reference, we also estimate the daily JGB zero curve from January 1999 through December 2011 using the method in Steeley (1991).[4]

II. Zero Curve Estimation Methods

In this section, we first define the zero yield and the zero curve, and summarize other basic items regarding interest rates required in this paper. Next, as a premise for considering the suitability of the estimation methods, we summarize the characteristics of undesirable zero curves. We then explain representative methods for the zero curve estimation and examine the characteristics of each method based on the two concepts of "degree of freedom" and "locality" from the perspectives of not estimating an undesirable zero curve and estimating a zero curve that accurately captures market prices.

A. Basic Items Concerning Interest Rates

We define the present time as t and the present value of bonds called discount bonds that will certainly pay cash flow 1 at the future time T as $Z(t, T)$. The zero yield $y(t, T)$ from t to T is defined as the yield to maturity of the discount bond.

$$y(t, T) = -\frac{1}{T-t}\log(Z(t, T)). \tag{1}$$

This paper estimates zero yields for all maturities, since we anticipate the potential use of zero yields for comparative analysis between yields with different maturities, as well as the use of the zero yield of a specific maturity. In other words, we estimate the curve connecting the zero yields with different maturities, which is called the zero curve. Specifically, we describe the zero curve at time t as a function of the remaining maturity x,

$$y(t, t+x), \tag{2}$$

and estimate the curve $y(t, t+x)$ for x at time t. In previous studies, for example, the discount rate $Z(t, t+x)$ is modeled in functional form as explained in Section II.D below. Hereafter $Z(t, t+x)$ is sometimes referred to as the discount function, as a function of x.

With the zero curve, it becomes possible to calculate the instantaneous spot interest rate, that is, the instantaneous interest rate at time t. The instantaneous spot interest rate $r(t)$ at time t is defined as equation (3).

4. The estimation data can be obtained from http://www.imes.boj.or.jp/research/papers/english/12-E-04.txt.

$$r(t) = \lim_{x \to 0} y(t, t + x). \tag{3}$$

Furthermore, with the zero curve it also becomes possible to calculate the implied forward rate. The implied forward rate is the interest rate over a period from a future point in time, defined at the present time. Specifically, the implied forward rate from time S to time T $(S < T)$ at t is defined so that the value derived when cash flow 1 at T is discounted through S by the implied forward rate and then discounted by the zero yield from t to S is equal to the present value when cash flow 1 at T is discounted by the zero yield from t to T. Therefore, the implied forward rate $f(t, S, T)$ from S to T at t is defined by equation (4).

$$Z(t, T) = \exp(-f(t, S, T)(T - S)) \exp(-y(t, S)(S - t)). \tag{4}$$

From equation (1) and equation (4), the implied forward rate $f(t, S, T)$ is calculated as equation (5) using the discount bond price at t.

$$f(t, S, T) = -\frac{1}{T - S} \log\left(\frac{Z(t, T)}{Z(t, S)}\right). \tag{5}$$

Additionally, the instantaneous forward rate at t for S as seen from $t, f(t, S)$, is defined as shown in equation (6).

$$\begin{aligned} f(t, S) &= \lim_{T \to S} f(t, S, T) \\ &= -\lim_{T \to S} \frac{1}{T - S} \log\left(\frac{Z(t, T)}{Z(t, S)}\right) = -\frac{\partial}{\partial S} \log(Z(t, S)). \end{aligned} \tag{6}$$

From equation (6) and $Z(S, S) = 1$, the following equation expresses the relation between the discount function $Z(t, t + x)$ and the instantaneous forward rate:

$$Z(t, t + x) = \exp\left(-\int_0^x f(t, t + s) ds\right). \tag{7}$$

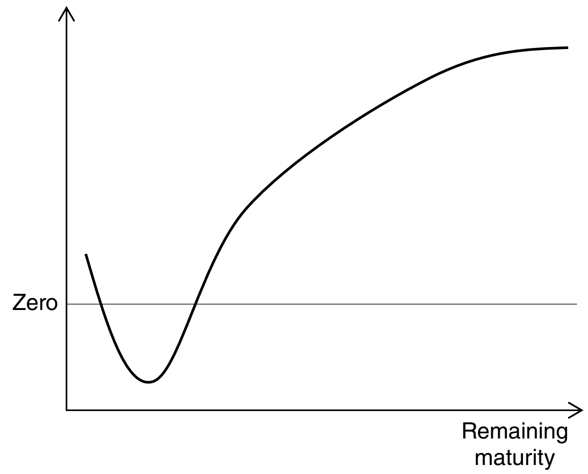
Furthermore, from equation (1) and equation (7), the relation between the zero yield and the instantaneous forward rate is as follows:

$$y(t, t + x) = \frac{1}{x} \int_0^x f(t, t + s) ds. \tag{8}$$

B. Undesirable Zero Curves

Estimation methods that can fully grasp the market prices of the bonds are preferable for estimating zero curves. If a zero curve is derived with an estimation method that does not adequately fit the market prices, there are concerns that the zero curve might

Figure 1 Zero Curve That Does Not Meet the Zero Interest Rate Constraint (Conceptual Diagram)



not fully reflect the information contained in the prices. Using such zero curves in interest rate analyses may lead to erroneous conclusions. On the other hand, methods selected based solely on the good fit to the market prices could result in an undesirable zero curve with improper interpolations. We now show the types of undesirable zero curves that may be estimated.

1. Violation of the zero interest rate constraint

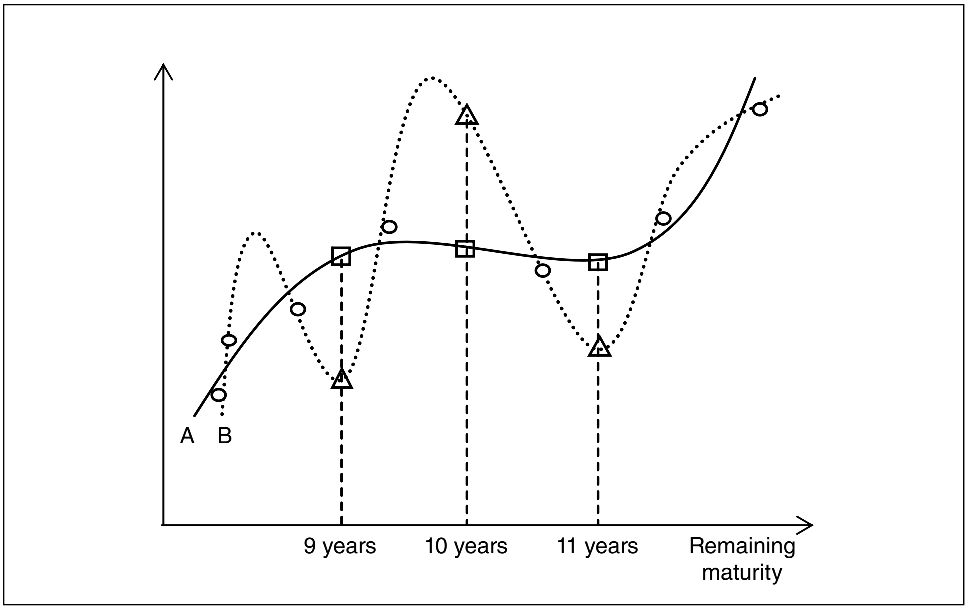
Because the estimated zero curve is the nominal interest rate, zero curves falling below zero for some maturities like the curve in Figure 1 are considered to be undesirable.

2. Excessive unevenness in the zero curve

Even for estimation methods with zero curves that do not go below zero, the zero curve may have an excessively uneven shape due to the characteristics of the estimation method.

Figure 2 is a conceptual diagram that shows the differences in the unevenness of the zero curves estimated using two estimation methods. In Figure 2, we assume that discount bonds with some maturities are traded on the market and zero curves are estimated based on these discount bond prices. Figure 2 shows that the two curves based on estimation method A and B both fit the market prices very well. However, method B has far greater unevenness. It is likely that curve B is unreasonable from the principle of the zero curve characteristic of little variation in zero yields with proximate maturities. The excessive unevenness of this type of zero curve, rather than reflecting the information contained in bond prices, may result from the peculiar characteristics of the estimation method. Thus, conducting analyses using a zero curve with very great unevenness like curve B could lead to erroneous results. For this reason, estimation methods with great unevenness in the zero curve like method B in Figure 2 cannot be considered desirable.

Figure 2 Unevenness of the Zero Curve (Conceptual Diagram)



Notes: ○ Zero yield calculated from discount bond market prices
□ Zero yield estimate calculated using estimation method A
△ Zero yield estimate calculated using estimation method B

3. Abnormal values

Some methods estimate zero yields with excessively high or excessively low values that are abnormal. Either these kinds of estimation methods have a poor fit with bond market prices due to their weak expressive power, or they have an over-fit with the market prices.

Figure 3 is a conceptual diagram presenting a zero curve with abnormal values. The curve in this figure overestimates the interest rate for some short-term maturities. This type of problem results from the characteristics of the estimation method and does not reflect the information contained in the bond prices.

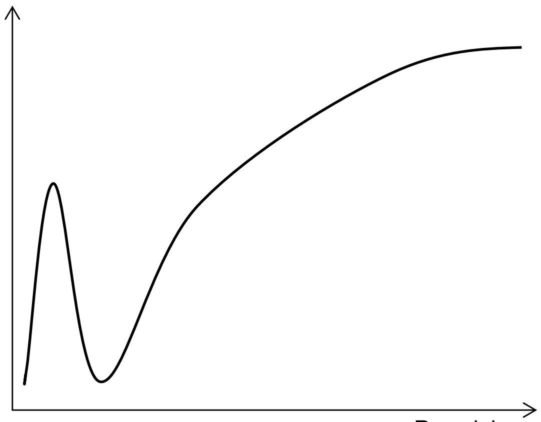
C. Definitions of Notations

Before explaining the contents of previous studies in Section II.D, we prepare the notations used in this paper.

For the sake of simplicity hereafter, the date when the zero curve is estimated is set at $t = 0$ and each point in time is expressed as the number of years from $t = 0$. The discount function $Z(0, x)$, the zero curve $y(0, x)$, and the instantaneous forward rate $f(0, x)$ are abbreviated as $Z(x)$, $y(x)$, and $f(x)$.

Next, we define the notations concerning fixed coupon-bearing bonds traded on the market. All bonds issued on or prior to the estimation date ($t = 0$) are expressed as $i(i \in \{1, \ldots, n_{name}\})$ with the notional amount of bond i expressed as N^i and the coupon rate (annual rate) as c^i. All the times when bond i generates cash flow are

Figure 3 Zero Curve with Abnormal Values (Conceptual Diagram)



expressed in terms of the number of years from $t = 0$, as $\mathbf{T}^i = \{T_1^i, \ldots, T_{n_{cf}^i}^i\}$. Here n_{cf}^i represents the number of times that bond i's cash flows are paid after $t = 0$. We assume that if $k < l$, then $T_k^i < T_l^i$. Additionally, $\mathbf{T}$ is defined as the union of the $\mathbf{T}^i$ of all the bonds issued at or before $t = 0$ as follows:

$$\mathbf{T} = \bigcup_{i=1}^{n_{name}} \mathbf{T}^i := \{T_1, \ldots, T_{n_{cf}}\},$$

$$T_j = \min_{\substack{i \in \{1,\ldots,n_{name}\} \\ 1 \le k \le n_{cf}^i}} \{T_k^i; T_k^i > T_{j-1}\},$$

$$T_1 = \min_{\substack{i \in \{1,\ldots,n_{name}\} \\ 1 \le k \le n_{cf}^i}} T_k^i.$$

We also define the notations regarding bonds traded on the market at the present time ($t = 0$). First, the bonds traded on the market at the present time are expressed as $\mathbf{I} = \{v_1, \ldots, v_{n_I}\}$. Then the present market price of each bond is expressed as $\mathbf{P} = (P^{v_1}, \ldots, P^{v_{n_I}})^{\mathrm{T}}$ where the price is the bare value, and the accrued interest at the execution of transactions is denoted by $\mathbf{A} = (A^{v_1}, \ldots, A^{v_{n_I}})^{\mathrm{T}}$.[5] The superscript T indicates a transposition of a vector or a matrix (here and hereafter, except as otherwise noted). We also define $\bar{\mathbf{P}}$ as $\bar{\mathbf{P}} = \mathbf{P} + \mathbf{A}$ and express it as $\bar{\mathbf{P}} = (\bar{P}^{v_1}, \ldots, \bar{P}^{v_{n_I}})^{\mathrm{T}}$ in the vector representation. For the details of the accrued interest calculation method, see Appendix 1, Section D.2.

5. We note that **P** and **A** are dependent on the present time. Here, for simplicity, the present time $t = 0$ is omitted. The model parameter $\boldsymbol{\alpha}$ introduced hereafter is also a variable dependent on the present time.

Finally, we prepare the notations used to express the theoretical price of bond v_i on the date when the zero curve is estimated. First, we define the vector $\bar{\mathbf{c}}^{v_i}$ concerning the cash flow of bond v_i as follows:[6]

$$\bar{\mathbf{c}}^{v_i} = \left(g(c^{v_i}, N^{v_i}, T_1), \ldots, g(c^{v_i}, N^{v_i}, T_j), \ldots, g(c^{v_i}, N^{v_i}, T_{n_{cf}})\right)^{\mathrm{T}},$$

$$g(c^{v_i}, N^{v_i}, T_j) = \begin{cases} \dfrac{c^{v_i} N^{v_i}}{2} & \text{if } T_j \in \mathbf{T}^{v_i}, T_j \neq T^{v_i}_{n^{v_i}_{cf}} \\ \dfrac{c^{v_i} N^{v_i}}{2} + N^{v_i} & \text{if } T_j = T^{v_i}_{n^{v_i}_{cf}} \\ 0 & \text{otherwise} \end{cases},$$

where $\bar{\mathbf{c}}^{v_i}$ is a vector of $n_{cf} \times 1$. If the j-th element of $\bar{\mathbf{c}}^{v_i}$ is expressed as $\bar{c}^{v_i}_j$, the theoretical price Q^{v_i} of bond v_i on the estimation date is as shown in equation (9).

$$Q^{v_i} = \sum_{j=1}^{n_{cf}} \bar{c}^{v_i}_j Z(T_j) - A^{v_i}. \tag{9}$$

Section II.D presents an outline of the representative previous studies on zero curve estimation. Previous studies all directly or indirectly model the discount function $Z(x)$ as a function of parameter $\boldsymbol{\alpha}$. To emphasize this point, we sometimes express the discount function $Z(x)$ as $Z(x; \boldsymbol{\alpha})$. Since the discount function depends on parameter $\boldsymbol{\alpha}$, the theoretical price of each bond also depends on parameter $\boldsymbol{\alpha}$ through equation (9). Thus, we express the theoretical price of bond v_i as $Q^{v_i}(\boldsymbol{\alpha})$ and write the theoretical prices of all bonds traded at $t = 0$ as the vector form $\mathbf{Q}(\boldsymbol{\alpha}) = (Q^{v_1}(\boldsymbol{\alpha}), \ldots, Q^{v_{n_I}}(\boldsymbol{\alpha}))^{\mathrm{T}}$.

D. Representative Estimation Methods Proposed in Previous Studies

In estimating the zero curve from the market prices of bonds, how we model the discount function $Z(x)$ is important. All previous studies that model the discount function $Z(x)$ adopt one of the following four methods: (1) the piecewise polynomial method; (2) the non-parametric method; (3) the polynomial function method; and (4) the method that assumes a specific functional form for the discount function. In this subsection, we present an outline of all the estimation methods that we deal with later in Section III.C.

In estimating the zero curve, in addition to modeling the discount function $Z(x)$, it is necessary to set the objective function to determine the parameters. In some previous studies, the objective function taking the weighted residual sum of squares is used, instead of the simple residual sum of squares of errors between the market prices and the theoretical prices based on the model.

For example, according to Bank for International Settlements (2005), the Bank of Canada and the Bank of Spain adopt the objective function taking the residual sum of squares weighted by the inverse of the bond durations. This objective function places an

6. The cash flow from fixed-coupon bonds issued since March 2001 is handled in this way, while a slightly different form is used for fixed-coupon bonds issued before March 2001. See Appendix 1, Section D.1 for the details.

emphasis on the fit to the prices of short-term bonds over long-term bonds. However, since we use fixed-coupon JGBs (two-year, five-year, 10-year, 20-year, and 30-year) for the zero curve estimation, there are more short-term bonds than long-term bonds. Therefore, using the residual sum of squares weighted by the squares of the inverse of the durations as the objective function might result in under-fitting to the market prices of long-term JGBs. Accordingly, we adopt the objective function with the simple residual sum of squares without taking any weights, as in McCulloch (1975), Steeley (1991), and many other previous studies. As Inui and Muromachi (2000) point out, adopting the simple residual sum of squares as the objective function may lead to heterogeneity of variance of residuals. Taking the weighted residual sum of squares of errors as the objective function is a solution to this problem. However, depending on the setting of weights, heterogeneity of variance of residuals does not always disappear; moreover, the estimation may result in a poor fit to a bond price. Considering these points, in this paper we assume the simple residual sum of squares as the objective function.

Furthermore, some of the previous studies adopt a function that adds a penalty term concerning the curvature of the instantaneous forward rate term structure to the residual sum of squares as the objective function, in order to estimate a zero curve with a smooth instantaneous forward rate term structure (Fisher, Nychka, and Zervos [1995], Waggoner [1997], Jarrow, Ruppert, and Yu [2004], etc.). This is called the smoothing spline method. While the zero curve estimated using such an objective function has a smoothed forward rate term structure, smoothing results in less fit to market prices. Additionally, in the smoothing method, there remains arbitrariness in the selection of criteria determining the level of smoothness. Therefore, we exclude the smoothing spline method from the several estimation methods analyzed in Section III. We select the models that estimate comparatively smooth zero curves without smoothing.

Considering the above, in this section we assume that the objective function is the simple residual sum of squares.

1. The piecewise polynomial method[7]

We here introduce previous studies that model the discount function by using the piecewise polynomial. As a means of modeling the discount function, most of the previous studies directly model the discount function or indirectly model the discount function by modeling the instantaneous forward rate.

First, we define the piecewise polynomial function. For this purpose, we must set the sequence of points known as knot points. The knot points are the following sequence:

$$u_m \leq u_{m+1} \leq \cdots \leq u_{n-1} \leq u_n,$$

where m and n are integers. When the knot points are given, for integer j, the piecewise polynomial function of degree l, $B(j, x)$ is continuous with respect to the real number x, and is the polynomial function on $[u_h, u_{h+1}]$ $(m \leq h \leq n-1)$, $(-\infty, u_m]$, and $[u_n, \infty)$.

7. Many previous studies model the discount function using piecewise polynomials known as spline functions. Hence, the piecewise polynomial method is also called the spline function method. The spline function of degree l is the piecewise polynomial whose derivatives from the first order to the $l-1$th order are all continuous.

a. Methods directly modeling the discount function

(1) The McCulloch (1975) method

McCulloch (1975) models the discount function $Z(x)$ as a linear combination of piecewise polynomial functions. First, in this method, the knot points are set as $0 = u_{-1} = u_0 = u_1 < u_2 < \cdots < u_{n_{knot}}$. Then McCulloch (1975) defines the piecewise polynomial $B(k, x)$ $(k = 0, \ldots, n_{knot})$ as the third-degree piecewise polynomial shown in equation (10).

For $k \neq n_{knot}$,

$$B(k,x) = \begin{cases} 0, & x \leq u_{k-1}, \\ \dfrac{(x-u_{k-1})^3}{6(u_k - u_{k-1})}, & u_{k-1} < x \leq u_k, \\ \dfrac{(u_k - u_{k-1})^2}{6} + \dfrac{(u_k - u_{k-1})(x - u_k)}{2} \\ \qquad + \dfrac{(x-u_k)^2}{2} - \dfrac{(x-u_k)^3}{6(u_{k+1} - u_k)}, & u_k < x \leq u_{k+1}, \\ (u_{k+1} - u_{k-1})(\dfrac{2u_{k+1} - u_k - u_{k-1}}{6} + \dfrac{x - u_{k+1}}{2}), & u_{k+1} < x, \end{cases}$$

For $k = n_{knot}$, $B(k, x) = x$.

$$\tag{10}$$

McCulloch (1975) expresses the discount function $Z(x)$ as a linear combination of piecewise polynomials defined in equation (10) as follows:

$$Z(x) = 1 + \sum_{k=0}^{n_{knot}} B(k,x)\alpha_k. \tag{11}$$

From the characteristics of the discount function, $Z(0) = 1$ must hold true. Modeling the discount function as in equation (11) is consistent with $Z(0) = 1$ because $B(k, 0) = 0$ $(k = 0, \ldots, n_{knot})$ from equation (10).

When we substitute equation (11) into equation (9), the theoretical price Q^{v_i} for bond v_i is expressed as a function of parameter $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_{n_{knot}})^{\mathrm{T}}$ as follows:

$$\begin{aligned} Q^{v_i}(\boldsymbol{\alpha}) &= \left(\sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i}\right) + \sum_{k=0}^{n_{knot}} \left(\sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i} B(k, T_j)\right) \alpha_k \\ &= \left(\sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i}\right) + (\bar{\mathbf{c}}^{v_i})^{\mathrm{T}} \mathbf{B} \boldsymbol{\alpha}, \end{aligned} \tag{12}$$

where $\mathbf{B}$ is the $n_{cf} \times (n_{knot} + 1)$ matrix with $B(k, T_j)$ as the (j, k)-th element. $\bar{\mathbf{c}}^{v_i}$ is the $n_{cf} \times 1$ vector with $\bar{c}_j^{v_i}$ as the j-th element.

As mentioned at the beginning of this section, the zero curve estimation in this paper uses the residual sum of squares of errors between bond market prices and theoretical prices as the objective function. In this case, the estimated value $\hat{\boldsymbol{\alpha}}$ of parameter $\boldsymbol{\alpha}$ is obtained as the solution to the following optimization problem:

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \left[\left(\tilde{\mathbf{P}} - \tilde{\mathbf{Q}}(\boldsymbol{\alpha})\right)^{\mathrm{T}} \left(\tilde{\mathbf{P}} - \tilde{\mathbf{Q}}(\boldsymbol{\alpha})\right) \right],$$

$$\tilde{\mathbf{P}} := (\bar{P}^{v_1} - \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_1}, \ldots, \bar{P}^{v_{n_I}} - \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_{n_I}})^{\mathrm{T}},$$

$$\tilde{\mathbf{Q}}(\boldsymbol{\alpha}) := (Q^{v_1}(\boldsymbol{\alpha}) - \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_1}, \ldots, Q^{v_{n_I}}(\boldsymbol{\alpha}) - \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_{n_I}})^{\mathrm{T}}. \tag{13}$$

Since $\tilde{\mathbf{Q}}(\boldsymbol{\alpha})$ is a linear function of $\boldsymbol{\alpha}$, equation (13) can be regarded as a least-squares optimization problem. Thus, the optimal solution $\hat{\boldsymbol{\alpha}}$ for parameter $\boldsymbol{\alpha}$ is obtained as equation (14),

$$\hat{\boldsymbol{\alpha}} = (\bar{\mathbf{c}}\mathbf{B}(\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}})^{-1}(\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}}\tilde{\mathbf{P}}, \tag{14}$$

where $\bar{\mathbf{c}}$ is set with $\bar{\mathbf{c}} = (\bar{\mathbf{c}}^{v_1}, \ldots, \bar{\mathbf{c}}^{v_{n_I}})^{\mathrm{T}}$ and it is the $n_I \times n_{cf}$ matrix. X^{-1} is the inverse matrix of the square matrix X.

Hereafter, this method is sometimes referred to as the McCulloch (1975) Method Modeling Discount Rates.

(2) The Steeley (1991) method

Steeley (1991), like McCulloch (1975), represents the discount function $Z(x)$ as a linear combination of piecewise polynomial functions. However, there is a difference in that Steeley (1991) models the discount function $Z(x)$ as shown in equation (15).

$$Z(x) = \sum_{k=-3}^{n_{knot}-1} B(k, x)\alpha_k. \tag{15}$$

Steeley (1991) also proposes a different functional form than McCulloch (1975) for the piecewise polynomial function $B(k, x)$. Steeley (1991) sets the knot points as $u_{-3} < \cdots < u_{n_{knot}} < u_{n_{knot}+1} < u_{n_{knot}+2} < u_{n_{knot}+3}$. Then $B(k, x)$ is recursively defined as shown in equation (16).

$$
\begin{aligned}
B(k,x) &= B_1(k,x) := \begin{cases} 1, & u_k \le x < u_{k+1} \\ 0 & \text{otherwise} \end{cases}, \\
B(k,x) &= B_D(k,x) = \frac{u_{D+k} - x}{u_{D+k} - u_{k+1}} B_{D-1}(k+1,x) \\
&\qquad + \frac{x - u_k}{u_{D+k-1} - u_k} B_{D-1}(k,x) \text{ for } D > 1.
\end{aligned} \tag{16}
$$

The function in equation (16) is called a B-spline function. Steeley (1991) proposes a model based on a piecewise cubic polynomial with $D = 4$ in equation (16). Here, we note that the following equation holds true from equation (15) and $Z(0) = 1$:

$$
\sum_{k=-3}^{n_{knot}-1} B(k,0)\alpha_k = 1. \tag{17}
$$

When we substitute equation (15) into equation (9), the theoretical price Q^{v_i} for bond v_i is represented as follows:

$$
\begin{aligned}
Q^{v_i}(\boldsymbol{\alpha}) &= \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i} \left(\sum_{k=-3}^{n_{knot}-1} B(k,T_j)\alpha_k \right) = \sum_{k=-3}^{n_{knot}-1} \left(\sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i} B(k,T_j) \right) \alpha_k \\
&= (\bar{\mathbf{c}}^{v_i})^{\mathrm{T}} \mathbf{B}\boldsymbol{\alpha},
\end{aligned} \tag{18}
$$

where $\mathbf{B}$ is the $n_{cf} \times (n_{knot} + 3)$ matrix with $B(k, T_j)$ as the (j,k)-th element.

From equation (17) and equation (18), parameter $\boldsymbol{\alpha}$ is estimated as the solution to the following constrained least-squares optimization problem:

$$
\begin{aligned}
\hat{\boldsymbol{\alpha}} &= \arg\min_{\boldsymbol{\alpha}} \left[\left(\bar{\mathbf{P}} - \mathbf{Q}(\boldsymbol{\alpha}) \right)^{\mathrm{T}} \left(\bar{\mathbf{P}} - \mathbf{Q}(\boldsymbol{\alpha}) \right) \right], \\
s.t. &\sum_{k=-3}^{n_{knot}-1} B(k,0)\alpha_k = 1.
\end{aligned} \tag{19}
$$

Solving this, the optimal solution $\hat{\boldsymbol{\alpha}}$ is as shown in equation (20),

$$
\hat{\boldsymbol{\alpha}} = \left((\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}} \bar{\mathbf{c}}\mathbf{B} \right)^{-1} (\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}} \bar{\mathbf{P}} + \frac{1 - \mathbf{B_0}^T \left((\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}} \bar{\mathbf{c}}\mathbf{B} \right)^{-1} (\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}} \bar{\mathbf{P}}}{\mathbf{B_0}^T \left((\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}} \bar{\mathbf{c}}\mathbf{B} \right)^{-1} \mathbf{B_0}} \left((\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}} \bar{\mathbf{c}}\mathbf{B} \right)^{-1} \mathbf{B_0}, \tag{20}
$$

where $\mathbf{B}_0 = (B(-3, 0), \ldots, B(n_{knot} - 1, 0))^{\mathrm{T}}$.

Hereafter, this method is sometimes referred to as the Steeley (1991) Method Modeling Discount Rates. The details of the estimation algorithm for this method are presented in Appendix 2.

b. Methods modeling instantaneous forward rates

(1) The Fisher, Nychka, and Zervos (1995) method

Fisher, Nychka, and Zervos (1995) represent the zero curve estimation methods using piecewise polynomial functions in the general form to model the discount function $Z(x)$, the zero yield $y(x)$, and the instantaneous forward rate $f(x)$, respectively. In this paper, among these we explain the case when $f(x)$ is modeled directly.[8]

Fisher, Nychka, and Zervos (1995) model the instantaneous forward rate $f(x)$ as a linear combination of piecewise polynomial functions as follows:

$$f(x) = \sum_{k=m}^{n} B(k, x)\alpha_k, \tag{21}$$

where $B(k, x)$ is a piecewise polynomial.

Using equation (7) and equation (21), the discount function can be expressed as follows:

$$\begin{aligned} Z(x) &= \exp\left(-\int_0^x f(s)ds\right) = \exp\left(-\int_0^x \sum_{k=m}^{n} B(k, s)\alpha_k ds\right) \\ &= \exp\left(-\sum_{k=m}^{n} (\int_0^x B(k, s)ds)\alpha_k\right) = \exp\left(-\sum_{k=m}^{n} \bar{B}(k, x)\alpha_k\right), \\ \bar{B}(k, x) &:= \int_0^x B(k, s)ds. \end{aligned} \tag{22}$$

When we substitute equation (22) into equation (9), the theoretical price Q^{v_i} can be written as follows:

$$\begin{aligned} Q^{v_i}(\boldsymbol{\alpha}) &= \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i} \exp\left(-\sum_{k=m}^{n} \bar{B}(k, T_j)\alpha_k\right) = (\bar{\mathbf{c}}^{v_i})^{\mathrm{T}} \exp(-\bar{\mathbf{B}}\boldsymbol{\alpha}), \\ \exp(-\bar{\mathbf{B}}\boldsymbol{\alpha}) &:= \left(\exp(-\sum_{k=m}^{n} \bar{B}(k, T_1)\alpha_k), \ldots, \exp(-\sum_{k=m}^{n} \bar{B}(k, T_{n_{cf}})\alpha_k)\right)^{\mathrm{T}}, \end{aligned} \tag{23}$$

8. The primary objective in Fisher, Nychka, and Zervos (1995) is to estimate a smooth instantaneous forward rate term structure, and their estimate is conducted with smoothing. In our paper, however, as stated at the beginning of this section, the goal is to estimate a zero curve that simultaneously achieves a good fit with market prices and an appropriate interpolation. Therefore, our paper does not place emphasis on making the instantaneous forward rate term structure smooth. In this respect, the Fisher, Nychka, and Zervos (1995) model explained here has a different objective function from their research. We advance our discussion using the simple sum of the squares of the residuals of the theoretical prices and the market prices as the objective function.

where $\bar{\mathbf{B}}_{j,k} = \bar{B}(k, T_j)$.

Because $Q^{v_i}(\boldsymbol{\alpha})$ is a nonlinear function of $\boldsymbol{\alpha}$, it is necessary to solve a nonlinear optimization problem in order to find the optimal solution for $\boldsymbol{\alpha}$. Accordingly, Fisher, Nychka, and Zervos (1995) simplify the optimization problem by conducting a first-order Taylor approximation around an arbitrary point $\boldsymbol{\alpha} = \boldsymbol{\alpha}^0$ in equation (23). Specifically, $Q^{v_i}(\boldsymbol{\alpha})$ is approximated as follows:[9]

$$
\begin{aligned}
Q^{v_i}(\boldsymbol{\alpha}) &\approx (\bar{\mathbf{c}}^{v_i})^{\mathrm{T}}\left(\exp(-\bar{\mathbf{B}}\boldsymbol{\alpha}^0) + \frac{\partial}{\partial\boldsymbol{\alpha}}\exp(-\bar{\mathbf{B}}\boldsymbol{\alpha})\Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^0}(\boldsymbol{\alpha}-\boldsymbol{\alpha}^0)\right),\\
&= Q^{v_i}(\boldsymbol{\alpha}^0) + (\bar{\mathbf{c}}^{v_i})^{\mathrm{T}}\frac{\partial}{\partial\boldsymbol{\alpha}}\exp(-\bar{\mathbf{B}}\boldsymbol{\alpha})\Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^0}(\boldsymbol{\alpha}-\boldsymbol{\alpha}^0),\\
&= Q^{v_i}(\boldsymbol{\alpha}^0) - (\bar{\mathbf{c}}^{v_i})^{\mathrm{T}}\left[\bar{\mathbf{B}} * \left(\exp(-\bar{\mathbf{B}}\boldsymbol{\alpha}^0)\mathbf{1}^{\mathrm{T}}\right)\right](\boldsymbol{\alpha}-\boldsymbol{\alpha}^0), \qquad (24)
\end{aligned}
$$

where $\mathbf{1}$ is the $(n-m+1)$ dimensional vector of $\mathbf{1} = (1, \ldots, 1)^{\mathrm{T}}$. The asterisk $*$ in equation (24) represents the element-by-element product of vectors or matrices. With the following equation (25):

$$
\begin{aligned}
\mathbf{X}^{v_i}(\boldsymbol{\alpha}^0) &= -(\bar{\mathbf{c}}^{v_i})^{\mathrm{T}}\left[\bar{\mathbf{B}} * \left(\exp(-\bar{\mathbf{B}}\boldsymbol{\alpha}^0)\mathbf{1}^{\mathrm{T}}\right)\right],\\
\mathbf{Y}^{v_i}(\boldsymbol{\alpha}^0) &= \bar{P}^{v_i} - Q^{v_i}(\boldsymbol{\alpha}^0) + \mathbf{X}^{v_i}(\boldsymbol{\alpha}^0)\boldsymbol{\alpha}^0, \qquad (25)
\end{aligned}
$$

the optimal solution $\hat{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^0)$ based on an approximation of equation (23) around an arbitrary point $\boldsymbol{\alpha} = \boldsymbol{\alpha}^0$ is the solution of the following optimization problem:

$$
\begin{aligned}
\hat{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^0) &= \arg\min_{\boldsymbol{\alpha}}\left[\left(\mathbf{Y}(\boldsymbol{\alpha}^0) - \mathbf{X}(\boldsymbol{\alpha}^0)\boldsymbol{\alpha}\right)^{\mathrm{T}}\left(\mathbf{Y}(\boldsymbol{\alpha}^0) - \mathbf{X}(\boldsymbol{\alpha}^0)\boldsymbol{\alpha}\right)\right],\\
\mathbf{X}(\boldsymbol{\alpha}^0) &:= (\mathbf{X}^{v_1}(\boldsymbol{\alpha}^0), \ldots, \mathbf{X}^{v_{n_I}}(\boldsymbol{\alpha}^0))^{\mathrm{T}},\\
\mathbf{Y}(\boldsymbol{\alpha}^0) &:= (\mathbf{Y}^{v_1}(\boldsymbol{\alpha}^0), \ldots, \mathbf{Y}^{v_{n_I}}(\boldsymbol{\alpha}^0))^{\mathrm{T}}. \qquad (26)
\end{aligned}
$$

There is no need to consider the constraint condition $Z(0) = 1$ on the discount function for the optimization problem in equation (26) because the following equation holds from equation (22):

$$
Z(0) = \exp\left(-\sum_{k=m}^{n}\bar{B}(k,0)\alpha_k\right) = \exp\left(-\sum_{k=m}^{n} 0 \times \alpha_k\right) = 1.
$$

Equation (26) can be solved as an unconstrained least-squares problem as follows:

$$
\hat{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^0) = \left(\mathbf{X}(\boldsymbol{\alpha}^0)^{\mathrm{T}}\mathbf{X}(\boldsymbol{\alpha}^0)\right)^{-1}\mathbf{X}(\boldsymbol{\alpha}^0)^{\mathrm{T}}\mathbf{Y}(\boldsymbol{\alpha}^0). \qquad (27)
$$

9. There are errors in the formula equivalent to equation (24) in the original paper.

However, this solution depends on the point where the Taylor approximation is conducted, $\boldsymbol{\alpha} = \boldsymbol{\alpha}^0$. Thus, Fisher, Nychka, and Zervos (1995) conduct a Taylor approximation of $Q^{v_i}(\boldsymbol{\alpha})$ around the point $\hat{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^0) \equiv \boldsymbol{\alpha}^1$ like that in equation (24), and calculate the optimal solution $\hat{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^1) \equiv \boldsymbol{\alpha}^2$ as in equation (27). They also repeat the same operation for $\boldsymbol{\alpha}^2$ and propose the convergence point of the optimal solution $\hat{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^i)$ as the optimal solution of the parameter.

While this is the estimation method proposed in Fisher, Nychka, and Zervos (1995), in implementing the estimation under this method, the piecewise polynomial function $B(k, x)$ in equation (21) must be determined. In the choice of estimation methods in Section III.C below, the methods considered for selection include (1) the method using the piecewise quadratic polynomial[10] proposed in McCulloch (1971) as $B(k, x)$ (hereafter, sometimes referred to as the McCulloch [1971] Method Modeling Instantaneous Forward Rates); and (2) the method using the B-spline function of degree two in Steeley (1991) as $B(k, x)$ (when $D = 3$ in equation [16]) (hereafter, sometimes referred to as the Steeley [1991] Method Modeling Instantaneous Forward Rates). As in the cases when the discount function is directly modeled in Section II.D.1.a.(1) and (2), the discount functions for both the two methods considered for selection are cubic functions.

2. The non-parametric method

a. The Tanggaard (1997) method

Unlike the above approaches, Tanggaard (1997) does not use a piecewise polynomial function to represent the discount function, but rather deals with the discount function of each term as a parameter.[11] In other words, Tanggaard (1997) represents the theoretical price $Q^{v_i}(\boldsymbol{\alpha})$ as follows:

$$Q^{v_i} = \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i} \alpha_j = (\bar{\mathbf{c}}^{v_i})^{\mathrm{T}} \boldsymbol{\alpha}, \quad \alpha_j = Z(T_j). \tag{28}$$

Because the theoretical price $Q^{v_i}(\boldsymbol{\alpha})$ is a linear function of $\boldsymbol{\alpha}$, the optimal solution $\hat{\boldsymbol{\alpha}}$ can be obtained by solving equation (29).

$$\min_{\boldsymbol{\alpha}} \left[(\bar{\mathbf{P}} - \mathbf{Q}(\boldsymbol{\alpha}))^{\mathrm{T}} (\bar{\mathbf{P}} - \mathbf{Q}(\boldsymbol{\alpha})) \right]. \tag{29}$$

Solving this, the optimal solution $\hat{\boldsymbol{\alpha}}$ is as shown in equation (30),

$$\hat{\boldsymbol{\alpha}} = (\bar{\mathbf{c}}^{\mathrm{T}} \bar{\mathbf{c}})^{-1} \bar{\mathbf{c}}^{\mathrm{T}} \bar{\mathbf{P}}, \tag{30}$$

where $\bar{\mathbf{c}} = (\bar{\mathbf{c}}^{v_1}, \ldots, \bar{\mathbf{c}}^{v_{n_I}})^{\mathrm{T}}$.

10. This is the differential function with respect to x of equation (10) (see McCulloch [1971]).

11. Aside from Tanggaard (1997) introduced here, other non-parametric methods include Carleton and Cooper (1976), and Houglet (1980).

3. The polynomial method

a. The Schaefer (1981) method

Schaefer (1981) represents the discount function $Z(x)$ using a linear combination of polynomials known as Bernstein polynomials. The Bernstein polynomial $B_D(k,x)$ of degree D is the polynomial defined as follows:

$$\begin{aligned} &B_D(k,x) = \sum_{j=0}^{D-k} (-1)^{j+1} \binom{D-k}{j} \frac{x^{k+j}}{k+j}, k > 0, \\ &\binom{D-k}{j} := \frac{(D-k)!}{(D-k-j)!j!}, \\ &B_D(0,x) = 1. \end{aligned} \tag{31}$$

Schaefer (1981) represents the discount function as a linear combination of the Bernstein polynomials as follows:

$$Z(x) = \sum_{k=0}^{D} B_D(k,x)\alpha_k.$$

Since $Z(0) = 1$ and $B_D(k,0) = 0$ for $k > 0$, we obtain $\alpha_0 = 1$ from the above equation. Therefore, $Z(x)$ is represented in the following form:

$$Z(x) = \sum_{k=0}^{D} B_D(k,x)\alpha_k = 1 + \sum_{k=1}^{D} B_D(k,x)\alpha_k. \tag{32}$$

When we substitute equation (32) into equation (9), the theoretical price Q^{v_i} of bond v_i is expressed as a function of the parameter $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_D)^T$ as follows:

$$\begin{aligned} Q^{v_i}(\boldsymbol{\alpha}) &= \left(\sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i} \right) + \sum_{k=1}^{D} \left(\sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i} B_D(k, T_j) \right) \alpha_k \\ &= \left(\sum_{j=1}^{n_{cf}} \bar{c}_j^{v_i} \right) + (\bar{\mathbf{c}}^{v_i})^{\mathrm{T}} \mathbf{B} \boldsymbol{\alpha}, \end{aligned} \tag{33}$$

where $\mathbf{B}$ is the $n_{cf} \times D$ matrix with $B_D(k,T_j)$ as the (j,k)-th element.

From equation (33), the optimal parameter $\hat{\boldsymbol{\alpha}}$ can be obtained by solving the following least-squares problem:

$$
\begin{aligned}
\hat{\boldsymbol{\alpha}} &= \arg\min_{\boldsymbol{\alpha}} \Big[\big(\tilde{\mathbf{P}} - \tilde{\mathbf{Q}}(\boldsymbol{\alpha})\big)^{\mathrm{T}} \big(\tilde{\mathbf{P}} - \tilde{\mathbf{Q}}(\boldsymbol{\alpha})\big) \Big], \\
\tilde{\mathbf{P}} &:= (\bar{P}^{v_1} - \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_1}, \ldots, \bar{P}^{v_{n_I}} - \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_{n_I}})^{\mathrm{T}}, \\
\tilde{\mathbf{Q}}(\boldsymbol{\alpha}) &:= (Q^{v_1}(\boldsymbol{\alpha}) - \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_1}, \ldots, Q^{v_{n_I}}(\boldsymbol{\alpha}) - \sum_{j=1}^{n_{cf}} \bar{c}_j^{v_{n_I}}).
\end{aligned} \tag{34}
$$

Thus, the optimal solution $\hat{\boldsymbol{\alpha}}$ becomes as shown in equation (35).

$$
\hat{\boldsymbol{\alpha}} = (\bar{\mathbf{c}}\mathbf{B}(\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}})^{-1}(\bar{\mathbf{c}}\mathbf{B})^{\mathrm{T}}\tilde{\mathbf{P}}. \tag{35}
$$

4. Parsimonious function methods

a. The Nelson and Siegel (1987) method

Nelson and Siegel (1987) use a parametric function in equation (36) to model the instantaneous forward rate $f(x)$.

$$
f(x) = \alpha_0 + \alpha_1 \exp\left(-\frac{x}{\alpha_3}\right) + \alpha_2 \frac{x}{\alpha_3} \exp\left(-\frac{x}{\alpha_3}\right). \tag{36}
$$

From equation (8) and equation (36), the zero curve $y(x)$ is calculated as follows:

$$
\begin{aligned}
y(x) = \frac{1}{x}\int_0^x f(s)ds &= \alpha_0 + \alpha_1 \left(\frac{1 - \exp(-x/\alpha_3)}{x/\alpha_3}\right) \\
&\quad + \alpha_2 \left(\frac{1 - \exp(-x/\alpha_3)}{x/\alpha_3} - \exp\left(-\frac{x}{\alpha_3}\right)\right).
\end{aligned} \tag{37}
$$

Here, the theoretical price Q^{v_i} of bond v_i is expressed by equation (9). Since the discount function $Z(x;\boldsymbol{\alpha})$ is a nonlinear function of the parameter $\boldsymbol{\alpha} := (\alpha_0, \alpha_1, \alpha_2, \alpha_3)^{\mathrm{T}}$ as shown in equation (37), the optimal solution $\hat{\boldsymbol{\alpha}}$ is obtained by solving the following nonlinear optimization problem:

$$
\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \Big\{ \big(\bar{\mathbf{P}} - \mathbf{Q}(\boldsymbol{\alpha})\big)^{\mathrm{T}} \big(\bar{\mathbf{P}} - \mathbf{Q}(\boldsymbol{\alpha})\big) \Big\}. \tag{38}
$$

b. The Svensson (1995) method

Svensson (1995) adds a new term to the functional form proposed in Nelson and Siegel (1987) (see equation [36]) for modeling the instantaneous forward rate to improve the expressive power of the instantaneous forward rate $f(x)$. Specifically, $f(x)$ is modeled using the following functional form:

$$f(x) = \alpha_0 + \alpha_1 \exp\left(-\frac{x}{\alpha_3}\right) + \alpha_2 \frac{x}{\alpha_3} \exp\left(-\frac{x}{\alpha_3}\right) + \alpha_4 \frac{x}{\alpha_5} \exp\left(-\frac{x}{\alpha_5}\right). \tag{39}$$

From equation (39), the zero yield $y(x)$ is calculated as follows:

$$\begin{aligned} y(x) &= \frac{1}{x}\int_0^x f(s)ds \\ &= \alpha_0 + \alpha_1\left(\frac{1-\exp(-x/\alpha_3)}{x/\alpha_3}\right) + \alpha_2\left(\frac{1-\exp(-x/\alpha_3)}{x/\alpha_3} - \exp\left(-\frac{x}{\alpha_3}\right)\right) \\ &\quad + \alpha_4\left(\frac{1-\exp(-x/\alpha_5)}{x/\alpha_5} - \exp\left(-\frac{x}{\alpha_5}\right)\right). \end{aligned} \tag{40}$$

From equation (40), as in the Nelson and Siegel (1987) method, the theoretical price Q^{v_i} of bond v_i is a nonlinear function of the parameter; therefore, the parameter estimation is solved using a nonlinear optimization problem.

E. Degree of Freedom and Locality of the Estimation Methods

In Section II.B, we presented undesirable zero curves. In this subsection, we introduce the two concepts, degree of freedom and locality, into the estimation methods to select methods that estimate desirable zero curves and reject ones which estimate undesirable zero curves. We then examine each of the zero curve estimation methods based on this perspective.

We begin with the explanation of the degree of freedom of the estimation method. In previous studies on the zero curve estimation methods introduced in Section II.D, the discount functions or the instantaneous forward rates are modeled with some functions. In this paper, we define the degree of freedom as the difference between the number of parameters of those functions and the number of constrained conditions imposed on these parameters. For example, the degree of freedom of the McCulloch (1975) Method Modeling Discount Rates explained in Section II.D.1.a.(1) is equal to the number of knot points + 1, while the degree of freedom of the Nelson and Siegel (1987) method in Section II.D.4.a is equal to four. In general, estimation methods with a low degree of freedom give a poor fit to the market prices, while methods with a high degree of freedom give a good fit to the prices. However, as shown in Section II.B, estimation methods with an excessively high degree of freedom may estimate zero curves with excessive unevenness or inappropriate interpolations. To avoid such an undesirable zero curve, it is necessary to use an estimation method without an excessively high degree of freedom.

Next, we explain the locality of the estimation method. For the sake of simplicity, we assume that the discount bonds are traded on the market and the zero curve is estimated from the discount bond prices. The concept of locality means the extent to which the shape of the estimated zero curve is changed when a discount bond price changes.

In other words, if a change in the price of a discount bond with a given maturity greatly changes the zero curve estimates with other maturities, the estimation method is said to have low locality. Conversely, if the price change of a discount bond does not greatly change the zero curve estimates for other maturities, the estimation method is said to have high locality. One advantage of estimation methods with high locality is that even when a bond with a remaining maturity has an abnormal price, this has almost no effect on the zero curve estimates for the other maturities. Another advantage of methods with high locality is that they are likely to estimate complex shaped zero curves with multiple inflection points.

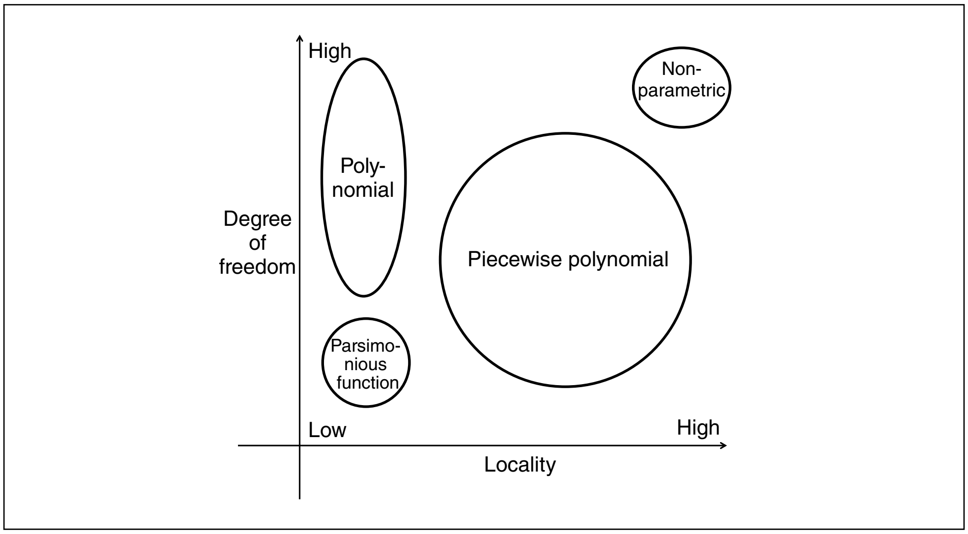
We now formulate the concept of locality described above. The original discount bond price data are denoted by $\mathbf{P}$, and $\mathbf{P} + \lambda P^{v_i}$ denotes the price data obtained when only the price of bond v_i with a remaining maturity $T^{v_i}_{n^{v_i}_{cf}}$ changes by λP^{v_i} while there are not any price changes in the other bonds. The zero curve estimated with method X from $\mathbf{P}$ is denoted by $\tilde{y}_X(x; \mathbf{P})$, and the zero curve estimated from $\mathbf{P} + \lambda P^{v_i}$ is denoted by $\tilde{y}_X(x; \mathbf{P} + \lambda P^{v_i})$. Then, the following index $l_X(\lambda, \varepsilon; T^{v_i}_{n^{v_i}_{cf}})$ defined in equation (41) may be considered as an index to measure the locality for the maturity $T^{v_i}_{n^{v_i}_{cf}}$ of method X:

$$\begin{aligned} l_X(\lambda, \varepsilon; T^{v_i}_{n^{v_i}_{cf}}) := & \int_0^{T^{v_i}_{n^{v_i}_{cf}} - \varepsilon} |\tilde{y}_X(x; \mathbf{P} + \lambda P^{v_i}) - \tilde{y}_X(x; \mathbf{P})|^2 dx \\ & + \int_{T^{v_i}_{n^{v_i}_{cf}} + \varepsilon}^{T_{n_{cf}}} |\tilde{y}_X(x; \mathbf{P} + \lambda P^{v_i}) - \tilde{y}_X(x; \mathbf{P})|^2 dx, \quad (41) \end{aligned}$$

where ε is a sufficiently small positive real number and λ is a real number. If the value calculated in equation (41) is small compared with those of the other estimation methods, at least, it is said that method X has higher locality around the remaining maturity $T^{v_i}_{n^{v_i}_{cf}}$ of bond v_i than the other methods. If the values calculated from equation (41) for all bonds (or all maturities) are small compared with those of the other estimation methods, then the method X has high locality for all maturities.

However, equation (41) cannot be considered an easy-to-use index because it depends on the choice of ε and λ. In addition, we must calculate indexes of all bonds to judge the locality for any maturity. Therefore, as an index to evaluate comparatively easily the locality of estimation methods, we now consider the ratio of the number of parameters that contribute to the determination of the discount rate with a remaining maturity to the degree of freedom. The lower this ratio is, the higher the locality. For example, the ratio for the Nelson and Siegel (1987) method for any maturity is equal to one. Next, we calculate the ratio for the Steeley (1991) Method Modeling Discount Rates in Section II.D.1.a.(2). If the knot points are set as $u_l = l (l = -3, \ldots, 33)$, the degree of freedom becomes 32. Using this, we calculate the ratios as follows. The ratio for maturities of less than a year is $4/32 = 0.125$, the ratio for maturities of 2–30 years in annual increments is $2/32 = 0.0625$, and the ratio for all other maturities

Figure 4 Characteristics of Each Method (Conceptual Diagram)



is $3/32 = 0.09375$. Hence, the Steeley (1991) Method Modeling Discount Rates has higher locality than the Nelson and Siegel (1987) method for all maturities.

Next, we see how the locality differs according to the kind of piecewise polynomial method. The locality of the Steeley (1991) Method Modeling Discount Rates is different from that of the McCulloch (1975) Method Modeling Discount Rates. Under the McCulloch (1975) Method Modeling Discount Rates, setting the knot points as $u_l = l(l = 0, \ldots, 30)$, $u_{-1} = 0$, the degree of freedom is equal to 31. The ratio between the number of parameters that contribute to the determination of the discount function and the degree of freedom is $2/31 = 0.0645$ for maturities of one year or less, $3/31 = 0.0968$ for maturities of two years or less, and $4/31 = 0.129$ for maturities of three years or less. In this way, it rises as the maturities increase. This shows that except for remaining maturities of one year or less, the Steeley (1991) Method Modeling Discount Rates has lower values than the McCulloch (1975) Method Modeling Discount Rates. We therefore conclude that the Steeley (1991) Method Modeling Discount Rates has higher locality.[12]

Figure 4 arranges the previous studies in Section II.D from the viewpoints of the concepts of degree of freedom and locality. The Tanggaard (1997) method, one of the non-parametric methods, directly estimates discount rates corresponding to times when all cash flows of all bonds are paid. Consequently, the degree of freedom of this method is much higher compared with the others. This method also has extremely high locality because it models the discount rate without assuming any specific functional form. In contrast, there are limits to the locality of the polynomial method and the parsimonious function method because they express the entire term structure of the discount function or the instantaneous forward rate as a single functional form. As for the degree

12. Like the McCulloch (1975) method, the McCulloch (1971) method also has lower locality compared with the Steeley (1991) method.

of freedom, we can deal with polynomial methods with various degrees of freedom by changing the degree of the polynomial used in the modeling. This is also true of the degree of freedom of the parsimonious function method. For the piecewise polynomial method, the degree of freedom depends on the number of knot points. The degree of freedom of this method rises as the number of knot points increases; as a result, the ratio of the number of parameters contributing to the discount rate of a given maturity to the degree of freedom decreases. This means that the locality of the piecewise method rises as the number of knot points increases. Thus the locality of the piecewise polynomial method is higher compared with the localities of the polynomial method and the parsimonious function method.

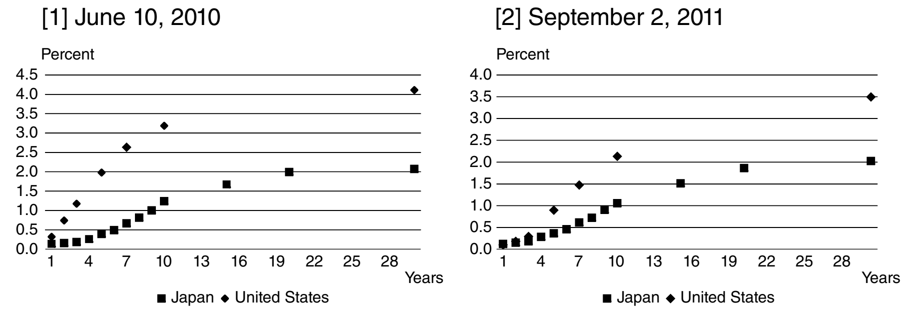
In the next section, we conduct comparative analyses on the eight zero curve estimation methods introduced in Section II.D. Here we briefly explain why we narrowed down the target to the eight estimation methods from among the many methods discussed in previous studies. The range of the previous studies on the non-parametric methods, the parsimonious function methods, and the polynomial methods is somewhat narrow. Hence, we have chosen one or two representative examples of each. On the other hand, various piecewise polynomial methods have been proposed in previous studies. We include the McCulloch (1971, 1975) methods and the Steeley (1991) method that have different locality in our analysis. Since some of the methods with piecewise polynomials model the discount rate as in Section II.D.1.a.(1) and (2) and others model the instantaneous forward rate as in Section II.D.1.b.(1), we analyze four piecewise polynomial methods based on the McCulloch (1971, 1975) methods and the Steeley (1991) method. Aside from these methods, other piecewise polynomial methods include Vasicek and Fong (1982) and McCulloch and Kochin (2000).[13] The localities of these methods are between or below those of the McCulloch (1971, 1975) methods and the Steeley (1991) method. Accordingly, we do not deal with any other piecewise polynomial methods in this paper's analysis aside from the four methods specified above.

III. Estimation Method Selection Criteria and Comparison Results

In this section, we select the estimation method that accurately grasps the characteristics of the term structure of JGB interest rates from the representative zero curve estimation methods introduced in Section II. To these ends, we summarize the JGB interest rate term structure characteristics in Section III.A. In Section III.B, we set criteria to exclude those zero curve estimation methods that generate undesirable zero curves as shown in Section II.B and to select a zero curve estimation method that grasps the JGB interest rate term structure characteristics. In Section III.C, we select the most appropriate zero curve method from the representative methods in light of our selection criteria through an analysis based on JGB market prices.

13. Vasicek and Fong (1982) conduct the variable change $x = 1 - \exp(-\alpha s)$ on the discount function $Z(x)$, and then use a piecewise polynomial to model the newly defined function $\tilde{Z}(s)(= Z(x))$. However, they do not propose the specific form of the piecewise polynomial used to model $\tilde{Z}(s)$. The McCulloch and Kochin (2000) method uses a piecewise polynomial to model the logarithmic function of the discount function.

Figure 5 Government Bond Interest Rate Term Structures of the United States and Japan



Notes: 1. Bloomberg ticker GJGBn Index for Japanese n-year interest rate (n =1, 2, . . . , 10, 15, 20, 30).
2. Bloomberg ticker USGGn Index for U.S. n-year interest rate of two years or more (n =1, 2, 3, 5, 7, 10, 30) and Bloomberg ticker USGG12M Index for one-year interest rate.

Source: Bloomberg.

A. Characteristics of the JGB Interest Rate Term Structure

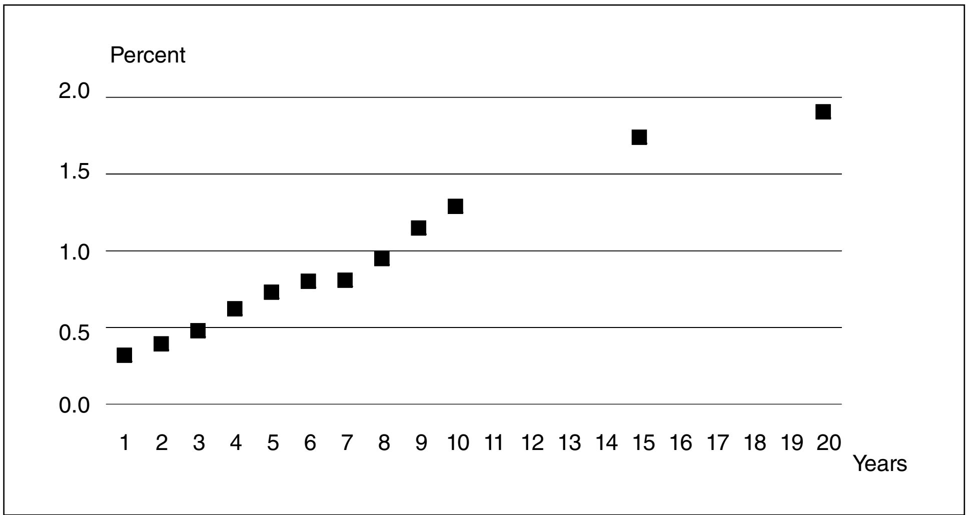
In this subsection, we note two representative characteristics of the term structure of JGB interest rates from the late 1990s until recently.

The first feature is that the interest rate term structure with remaining maturities of up to around three years shows a flat curve near zero. Especially during the quantitative easing policy period from 2001 to 2006 and the global financial crisis since the summer of 2007, the JGB interest rate term structure generally had this kind of shape. Figure 5 compares the term structures of U.S. and Japanese government bond interest rates since the financial crisis. During this period, both countries implemented *de facto* zero interest rate policies, but at certain points in time the shapes of the term structures for short-term maturities differed. In Figure 5 [1], the slope of the U.S. government bond yield curve increases from remaining maturities of two years, while the Japanese curve remains near zero through remaining maturities of around three years.

When estimating the zero curve from JGB market prices, an undesirable method may fail to accurately grasp the curve characteristics introduced above. For example, a portion of the estimated zero curve for short maturities might fall below zero. As shown in Section II.D, the polynomial and parsimonious function methods express the entire zero curve as a polynomial or a specific function. Therefore, it may be difficult to fully capture the above-mentioned characteristics of the JGB term structure due to the low locality of these methods.

In recent years, not only the JGB interest rate term structure but also government bond interest rate term structures in the United States and Europe have been flattening near zero for short-term maturities. Looking at the U.S. interest rate term structure in Figure 5 [2], compared with Figure 5 [1], the term structure is turning flatter near zero through maturities up to around three years. For this reason, at present, estimations

Figure 6 JGB Interest Rate Term Structure (February 17, 2009)



Note: n-year interest rate indicates Bloomberg ticker GJGBn Index.

Source: Bloomberg.

with low locality may not have sufficient expressiveness to estimate the zero curves in the United States and European countries.[14]

The second feature frequently seen in the JGB interest rate term structure is that it has a complex shape with multiple inflection points. For example, as in Figure 6, the seven-year interest rate sometimes becomes relatively low compared with the six-year and eight-year rates.[15]

Zero curve estimation methods with a low degree of freedom cannot capture curves with this kind of complex shape. When the degree of freedom of the estimation method is too high, however, it may estimate a zero curve with excessive unevenness as seen in Section II.B. Consequently, we need an estimation method with a suitable degree of freedom that is neither too low nor too high, so that it can grasp such characteristics of the JGB interest rate term structure. In addition, an estimation method that has a large influence on the estimates for other maturities to capture such uneven shapes is undesirable. Therefore, we believe that the zero curve estimation method must have high locality to grasp the characteristics of the JGB interest rate term structure, which has a complex shape with multiple inflection points.

14. In the United States, the Federal Reserve Board's (FRB's) economists estimate the zero curve based on the method in Svensson (1995) (see Gürkaynak, Sack, and Wright [2007] for the details of the estimation method). Our calculations of the zero yield for short-term maturities of one or two months with the zero curve estimation parameters released on the FRB homepage frequently resulted in estimates below zero during periods when the yield curve became steep under the low interest rate environment after the middle of 2009.
15. The maturity of the cheapest-to-deliver of the JGB futures is around seven years. Thus, the prices of JGBs with remaining maturities of around seven years have often been influenced by the prices of JGB futures. In particular, when the prices of JGB futures rose substantially with the flight to quality in the second half of 2008, JGBs with remaining maturities of around seven years were traded at a higher price than those with remaining maturities of six and eight years. In other words, the seven-year interest rate was lower compared with interest rates of other maturities.

B. Estimation Method Selection Criteria

In this subsection, we set criteria to exclude methods that estimate undesirable zero curves as shown in Section II.B and select methods that capture features of the JGB interest rate term structure shown in Section II.A above.

Criterion (1): Zero curve estimates should not fall below zero

As shown in Section III.A, the JGB interest rate term structure often takes a flat shape near zero for short-term maturities. Some estimation methods cannot capture this shape, and may estimate the zero curve below zero for some maturities. As explained in Section II.B, zero curves that fall below zero for some maturities are undesirable. Accordingly, we exclude estimation methods that estimate a comparatively large number of zero yield estimates for remaining maturities of 0.5, 1, 1.5, and 2 years falling below zero during the estimation period.

Criterion (2): Zero curve estimates should not include abnormal values

As shown in Section II.B, some estimation methods may estimate the zero curve with abnormal values such as extremely high or extremely low estimates. These methods end up under-fitting JGB market prices due to their low degree of freedom, or over-fitting due to their high degree of freedom. We estimate the zero yield with a specific maturity using each estimation method on each estimation date; in addition, we calculate the standard deviation of those zero yield estimates. Based on this, we regard estimates outside the range of ± 2 standard deviations as abnormal values. We then reject estimation methods with a relatively high frequency of abnormal values. This selection criterion excludes estimation methods with degrees of freedom that are either too high or too low.

Criterion (3): Theoretical prices should have a good fit with market prices

As noted in Section III.A, the JGB interest rate term structure often has a complex shape with multiple inflection points. We judge the extent of the fit of the theoretical prices to the market prices by whether this complex shape is accurately grasped. Specifically, we evaluate it using the residual sum of squares of errors of the market prices and the theoretical prices on the estimation date. This criterion is used to reject estimation methods that have poor fits to the market prices. While it does not reject estimation methods that have over-fits to the market prices, these methods can be rejected by criteria (2) above and (4) below.

Criterion (4): The shape of curve should not be extremely uneven

As stated in Section II.B, it is undesirable for the zero curve to be extremely uneven. If the zero yield estimates include abnormal values, there may be large rises and falls in the zero curve. These are rejected by criterion (2). Estimation methods that are not rejected by criterion (2), however, may include methods which estimate zero curves with extreme unevenness. Thus, to exclude this possibility, we check whether or not the curvature calculated using the zero yield estimates for specific maturities ranging from 0.5 to 20 years in half-year increments from equation (42) shows excessive values.

$$\sum_{j=2}^{39} (y(0.5(j+1)) - 2y(0.5j) + y(0.5(j-1)))^2. \tag{42}$$

In this paper, we exclude those estimation methods that frequently generate undesirable zero curves in light of criteria (1) and (2) from the selection candidates. The selection results are presented in Section III.C.3. Next, the most desirable estimation method is selected based on the perspectives in (3) and (4) from among those estimation methods that are not excluded in the above process.

C. Comparative Analysis of Estimation Methods

In this subsection, we select the most suitable estimation method for JGB price data from the representative zero curve estimation methods in Section II in light of the criteria presented in Section III.B.

1. Outline of the JGB price data used in the estimations

In our comparative analysis, we use price data on fixed coupon-bearing JGBs (two-year, five-year, 10-year, 20-year, and 30-year bonds). The price data are Japan Bond Trading Co., Ltd. JGB closing prices obtained from the NEEDS provided by Nikkei Digital Media Inc. The zero curve estimation period covers all business days from January 4, 1999 through December 30, 2010.

Figure 7 presents the number of fixed coupon-bearing JGB issues over the estimation period. As shown in the figure, the total number of issues was around 150 in January 1999, while it has grown to about 300 in recent years. This is a result of increased issuance of five-year, 20-year, and 30-year bonds.

In this way, the JGB issuance conditions have changed since the 2000s. In previous studies and surveys on the estimation of the JGB zero curve (Komine *et al*. [1989], Oda [1996], Inui and Muromachi [2000], and Kawasaki and Ando [2005]), price data before 2000 are used. In contrast, the estimation results in this paper take into consideration the change in issuance conditions from the 2000s described above.

2. Estimation methods for our analysis

The estimation methods for our analysis are the eight methods introduced in Section II.D. In this subsection, we present supplementary notes on each method.

(1) The McCulloch (1975) Method Modeling Discount Rates in Section II.D.1.a.(1)
The knot points are set in values incremented by one from 0 to 30.

(2) The Steeley (1991) Method Modeling Discount Rates in Section II.D.1.a.(2)
The knot points are set in values incremented by one from −3 to 33.

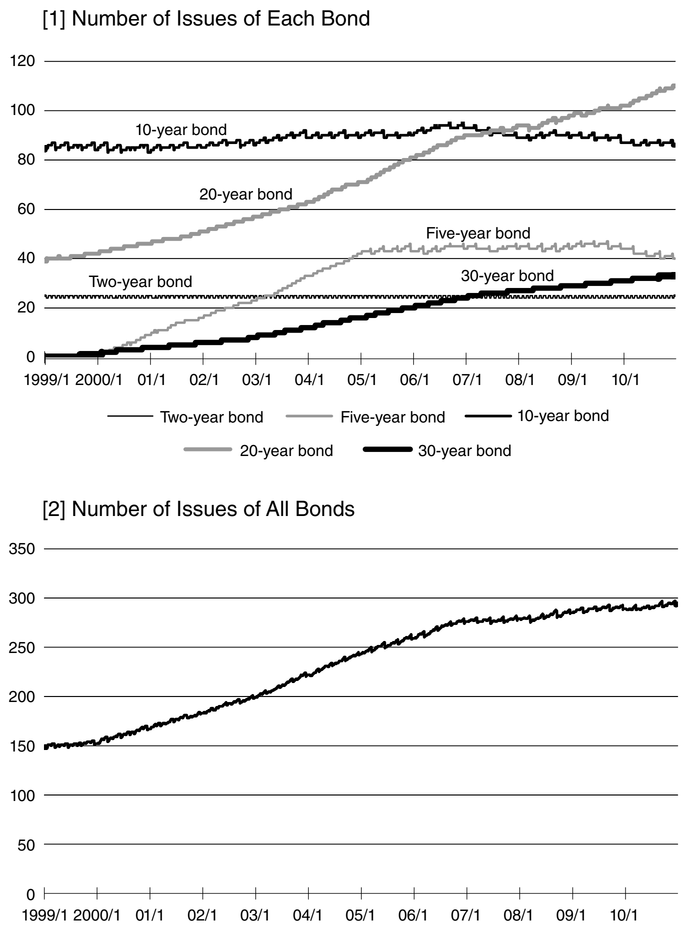
(3) The McCulloch (1971) Method Modeling Instantaneous Forward Rates in Section II.D.1.b.(1)
The knot points are set in values incremented by one from 0 to 30.

(4) The Steeley (1991) Method Modeling Instantaneous Forward Rates in Section II.D.1.b.(1)
The knot points are set in values incremented by one from −2 to 32.

(5) The Tanggaard (1997) method in Section II.D.2.a
The discount rates corresponding to times until cash flows are not paid cannot be directly estimated from this model. Hence, this paper estimates these yields by linear interpolation.

(6) The Schaefer (1981) method in Section II.D.3.a
The Bernstein polynomial is of the fifth degree.

Figure 7 Number of Issues



(7) The Nelson and Siegel (1987) method in Section II.D.4.a
The Nelder-Mead method is used in the parameter estimation.
(8) The Svensson (1995) method in Section II.D.4.b
This is the same as (7).

3. Comparison results: Exclusion of undesirable estimation methods

We now summarize the results of comparing the eight methods in light of the criteria (1) and (2) presented in Section III.B.

a. Number of times that zero yield estimates fall below zero

Table 1 shows the number of times that the zero yield estimates for remaining maturities of 0.5, 1, 1.5, and 2 years drops below zero during the estimation period under each method. The table shows that while the piecewise polynomial methods do not give any estimates below zero, the other methods do give estimates below zero. In particular, the polynomial method and the parsimonious function methods generate far larger numbers of estimates below zero compared with the other methods.

Table 1 Number of Times That Interest Rate Estimates Are Less than Zero

Piecewise polynomial methods			
Methods modeling discount rates		Methods modeling instantaneous forward rates	
McCulloch (1975)	Steeley (1991)	McCulloch (1971)	Steeley (1991)
0	0	0	0

Polynomial method	Non-parametric method	Parsimonious function methods	
Schaefer (1981)	Tanggaard (1997)	Nelson and Siegel (1987)	Svensson (1995)
1,084	30	2,162	619

Note: Total of 11,788 samples (4 samples/day × 2,947 days).

Figure 8 Examples of Zero Yield Estimates Below Zero

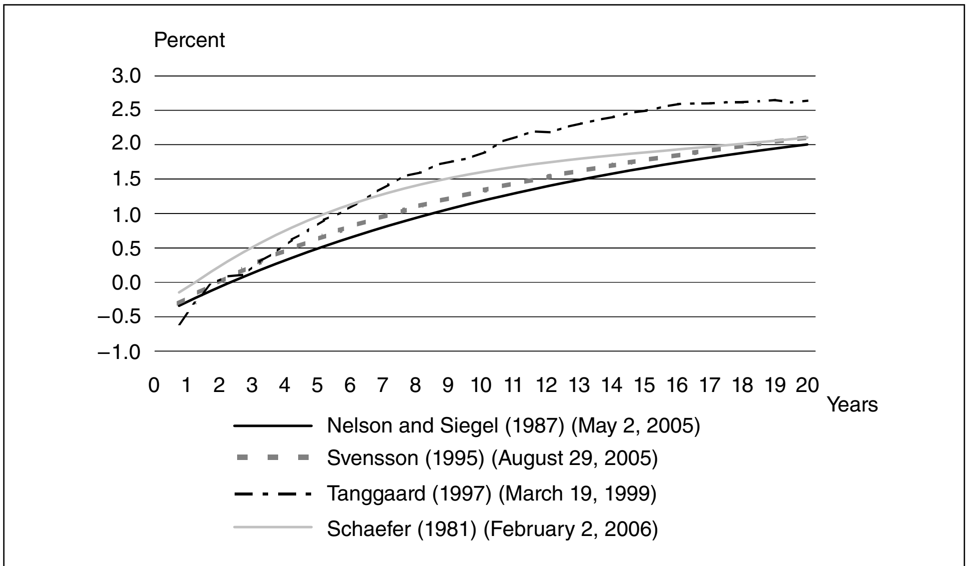


Figure 8 presents examples of zero curves on dates when estimates actually fell below zero for estimation methods with many estimates falling below zero. As shown in the figure, the polynomial method and the parsimonious function methods generate frequent estimation results with the zero yields below zero for certain maturities during the quantitative easing policy period. This is considered to result from the low locality of these methods as explained in Section III.A. Estimates below zero were also observed under the non-parametric method.

Given these results, we conclude that the polynomial method (the Schaefer [1981] method), the parsimonious function methods (the Nelson and Siegel [1987] method and the Svensson [1995] method), and the non-parametric method (the Tanggaard [1997] method) are not appropriate as methods to capture the characteristics of the JGB interest rate term structure.

Table 2 Number of Abnormalities

	Piecewise polynomial methods				Polynomial method	Non-parametric method	Parsimonious function methods	
	Methods modeling discount rates		Methods modeling instantaneous forward rates					
Remaining maturity	McCulloch (1975)	Steeley (1991)	McCulloch (1971)	Steeley (1991)	Schaefer (1981)	Tanggaard (1997)	Nelson and Siegel (1987)	Svensson (1995)
1 year	0	0	2	2	36	11	0	88
2 years	0	0	0	0	0	2	5	0
3 years	0	0	0	0	0	0	101	0
4 years	0	0	0	0	0	0	15	0
5 years	0	0	0	0	0	0	0	0
6 years	0	0	0	0	0	0	0	0
7 years	0	0	0	0	0	0	0	0
8 years	0	0	0	0	0	0	0	0
9 years	0	0	0	0	0	0	0	0
10 years	0	0	0	0	0	0	0	0
11 years	0	0	0	0	0	0	0	0
12 years	0	0	0	0	0	0	0	0
13 years	0	0	0	0	0	0	0	0
14 years	0	0	0	0	0	0	0	0
15 years	0	0	0	0	0	0	0	0
16 years	0	0	0	0	0	0	0	0
17 years	0	0	0	0	0	0	0	0
18 years	0	0	0	0	0	0	0	0
19 years	0	0	0	0	0	0	0	0
20 years	0	0	0	0	0	0	0	0

Notes: 1. There are 2,947 samples for each item (January 4, 1999 to December 30, 2010).
2. The mean and the standard deviation are calculated from the estimates generated by the eight estimation methods at each estimation date, and estimates outside the range of ± 2 standard deviations of the mean are defined as abnormal values.

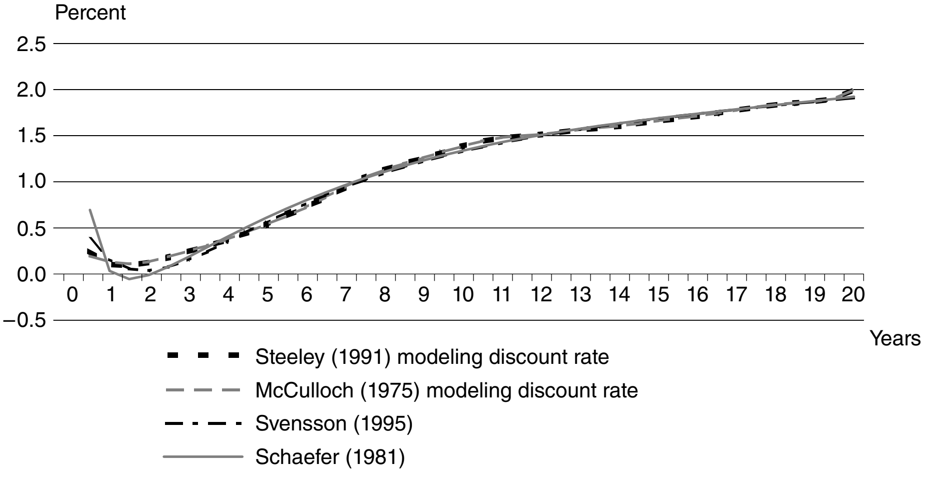
b. Number of times that the zero yield estimates take abnormal values

Table 2 summarizes the number of times that the zero yield estimates for each maturity take abnormal values during the estimation period under each method.

The table shows that for short-term maturities, the polynomial method (the Schaefer [1981] method) and the parsimonious function methods (the Nelson and Siegel [1987] method and the Svensson [1995] method) give more abnormal values compared with the other estimation methods. Abnormal values under these methods are caused as a result of estimates below zero or estimates with far larger values than those of other methods.

Figure 9 is an example of when the estimates for short-term maturities based on the polynomial methods and the parsimonious methods are higher than those of other

Figure 9 Examples of Zero Yield Estimates with Abnormal Values



methods. For a remaining maturity of 0.5 years, the zero yield estimate based on the Schaefer (1981) method is higher than the estimates based on the other methods. The reason the zero yield estimates for short-term maturities under the polynomial method and the parsimonious function methods take abnormal values is that these methods cannot capture the flat shape of the JGB yield curve near zero for short-term maturities often seen during the period of quantitative easing policy due to their low locality.

In light of whether the zero yield estimates take abnormal values, we conclude that the polynomial method and the parsimonious function methods are not appropriate for JGB zero curve estimates.

Given the above results, we conclude that aside from the McCulloch (1975) Method Modeling Discount Rates, the Steeley (1991) Method Modeling Discount Rates, the McCulloch (1971) Method Modeling Instantaneous Forward Rates, and the Steeley (1991) Method Modeling Instantaneous Forward Rates, the other estimation methods do not meet the selection criteria. Thus, we exclude methods except the above four methods from our analysis.

4. Comparison results: Selection of the optimal estimation method

In this subsection, we select the most appropriate estimation method from the four remaining methods (the McCulloch [1975] Method Modeling Discount Rates, the Steeley [1991] Method Modeling Discount Rates, the McCulloch [1971] Method Modeling Instantaneous Forward Rates, and the Steeley [1991] Method Modeling Instantaneous Forward Rates) from the perspectives of the fit of theoretical prices to market prices and the low unevenness of the zero curve.

a. Fit of theoretical prices to market prices

Table 3 compares the residual sum of squares under the four estimation methods during the estimation period. The table indicates that the residual sum of squares of the Steeley (1991) Method Modeling Discount Rates has the lowest mean value, standard deviation, and maximum value among the four methods during the estimation period.

Table 3 Comparison Results for Fitness to Market Prices

	Piecewise polynomial methods			
	Methods modeling discount rates		Methods modeling instantaneous forward rates	
	McCulloch (1975)	Steeley (1991)	McCulloch (1971)	Steeley (1991)
Maximum	29.178	27.058	29.126	29.116
Minimum	0.185	0.144	0.198	0.182
Mean	3.414	3.093	3.598	3.807
Standard deviation	4.728	4.441	4.714	5.041

Notes: 1. Mean, standard deviation, maximum, and minimum denote the mean value, standard deviation, maximum value, and minimum value of the residual sum of squares on each estimation date.
2. The unit is the square of yen.

Table 4 Comparison of Curvatures

Piecewise polynomial functions			
Methods modeling discount rates		Methods modeling instantaneous forward rates	
McCulloch (1975)	Steeley (1991)	McCulloch (1971)	Steeley (1991)
1.587×10^{-2}	1.470×10^{-2}	1.930×10^{-2}	1.991×10^{-2}

Notes: 1. The above figures show the mean values of the curvatures (equation [42]) on each estimation date during the estimation period.
2. The unit is the square of percent.

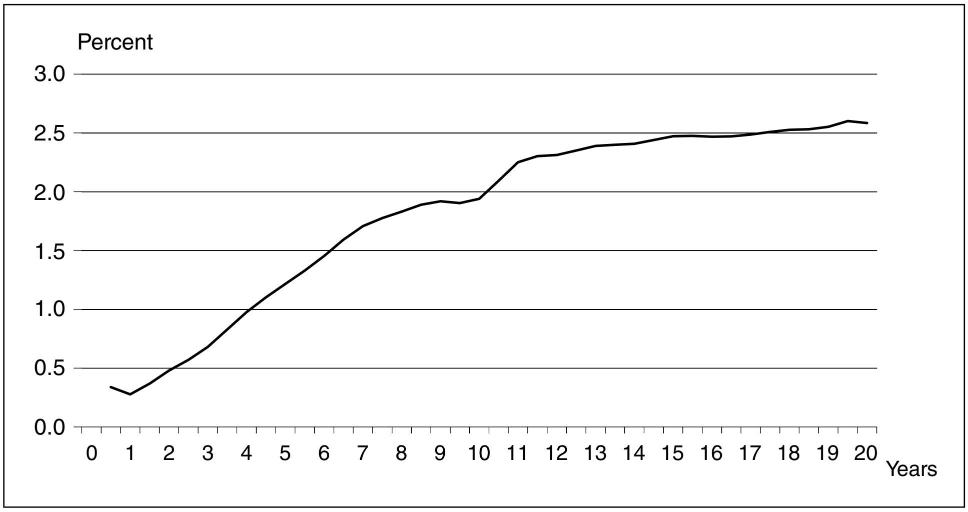
It appears that this result occurs because the Steeley (1991) Method Modeling Discount Rates has the highest locality among these four methods.

b. Unevenness of zero curves

Table 4 summarizes the averaged value of the curvature measuring the extent of the unevenness of the zero curve during the estimation period. While the table does not show great differences among the methods, the curve estimated based on the Steeley (1991) Method Modeling Discount Rates has the lowest curvature.

Figure 10 shows the estimation results on the date (February 8, 2000) when the zero curve under the Steeley (1991) Method Modeling Discount Rates has its highest curvature during the estimation period. In this figure, the zero yield with a maturity of 9.5 years is lower than those with maturities of both nine years and 10 years. This is considered to lead to the increase in the curvature of the zero curve. This reflects the fact that the price of the JGB with a remaining maturity of about 9.5 years was actually higher than the prices of the nine- and 10-year JGBs; hence, the interpolation is deemed to be appropriate. Considering this, we conclude that the unevenness of zero curves under the Steeley (1991) Method Modeling Discount Rates is appropriate.

Figure 10 Zero Curve Using the Steeley (1991) Method Modeling Discount Rates



Note: Zero curve estimated on February 8, 2000, when the zero curve under the Steeley (1991) Method Modeling Discount Rates had its highest curvature during the estimation period.

Summarizing all the above results, among the eight methods considered in our analysis, the Steeley (1991) Method Modeling Discount Rates is judged the most appropriate for estimating the JGB market zero curve. In the next section, we clarify the characteristics of this method through comparisons with the other estimation methods.

IV. Characteristics of the Steeley (1991) Method Modeling Discount Rates

In Section III, the Steeley (1991) Method Modeling Discount Rates (hereafter, the Steeley model) was selected as the optimal estimation method in light of the criteria set in this paper. In this section, we first clarify the characteristics of the Steeley model in Section IV.A by comparing the Steeley model with the Nelson and Siegel (1987) method (hereafter, the NS model). Specifically, we compare the two models' results in terms of (1) the shape of curves, (2) the short-term interest rates, and (3) the long-term interest rates. Next, in Section IV.B, we consider the curve based on the smoothing spline method and compare it with the curve based on the Steeley model.

A. Comparison of the Steeley Model with the NS Model

1. Comparison of zero curve shapes

Figure 11 presents the zero curves on December 30, 2010 estimated based on the Steeley model and the NS model. In Section III.A, we pointed out the flat shape of the curve near zero for short-term maturities as one of the features of the JGB interest rate term structure. The curve estimated based on the Steeley model takes this type of shape.

Figure 12 presents the comparison of the two models' curves on August 20, 2008. Around this time, against the backdrop of a rise in the price of JGB futures, the prices

Figure 11 Comparison of Curve Shapes (December 30, 2010)

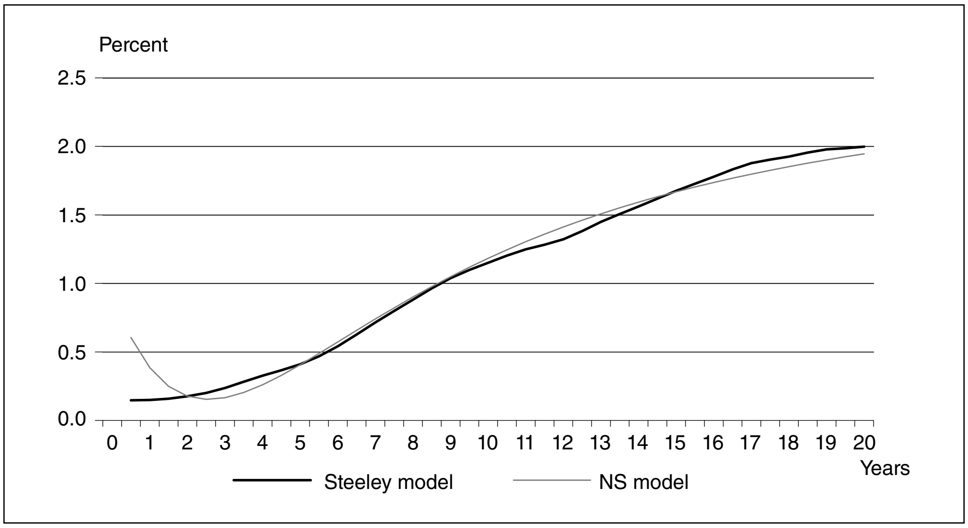
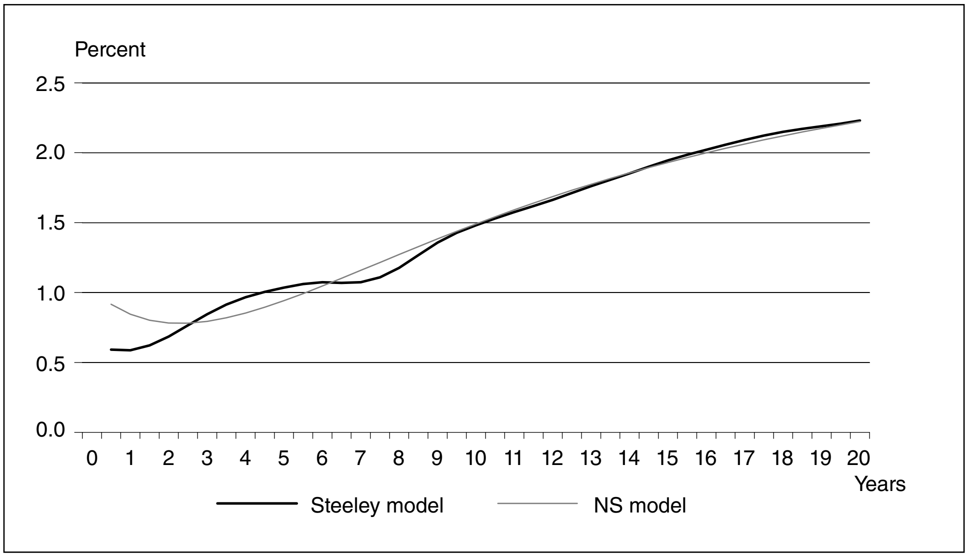


Figure 12 Comparison of Curve Shapes (August 20, 2008)

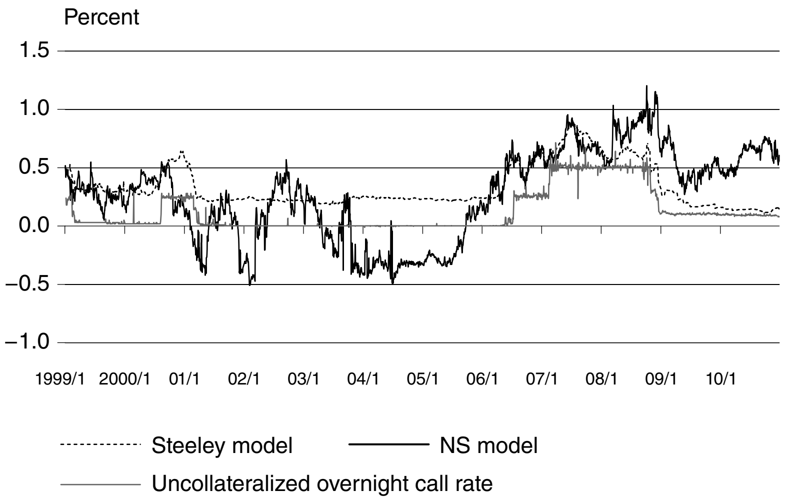


of JGBs with remaining maturities of around seven years rose relative to those of other JGBs. This means that the seven-year interest rate was low compared with the six- and eight-year interest rates. In fact, the zero curve estimated by the Steeley model captures this kind of yield curve shape. In contrast, the zero curve estimated by the NS model does not capture this unevenness from the six-year through the eight-year interest rates. It seems that the Steeley model is able to capture this shape because of its high locality.

2. Comparison of short-term interest rates

Figure 13 presents the short-term interest rate estimates using the Steeley model and the NS model.

Figure 13 Comparison of Short-Term Interest Rates (Six-Month Yields)



Looking at the figure, the six-month zero yield estimates under the NS model frequently fall below zero during the quantitative easing policy period. In contrast, the six-month zero yield estimates during this period based on the Steeley model remain steady near zero, without falling below zero.[16]

During periods other than the quantitative easing policy period as well, estimates based on the Steeley model remain steady without widely diverging from the movements of the uncollateralized overnight call rate, unlike the estimates using the NS model. In this way, the Steeley model is deemed to steadily estimate short-term interest rates, especially short-term interest rates during monetary easing periods in Japan.

3. Comparison of long-term interest rates

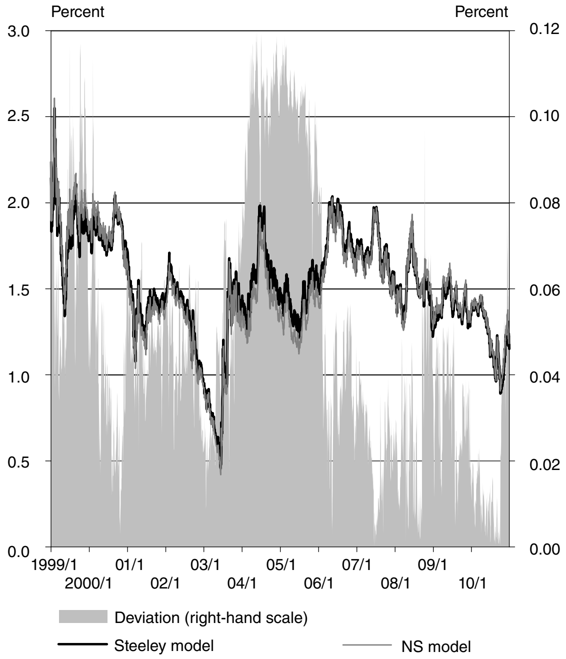
Figure 14 presents the long-term interest rate estimates (10-year zero yield and 20-year zero yield) using the Steeley model and the NS model.

Comparing the estimates using the two models, we find that the largest divergence is only around 10 basis points for the 10-year zero yield. For the 20-year zero yield as well, the largest divergence is also around 10 basis points. This suggests that there is no substantial difference between the two models in the estimates of long-term interest rates. However, the selection of the estimation method is important for analyses using multiple estimates on the curve, such as analyses using the spread between the short-term and long-term interest rates, because the short-term interest rate estimates differ depending on the choice of models as seen in Section IV.A.2.

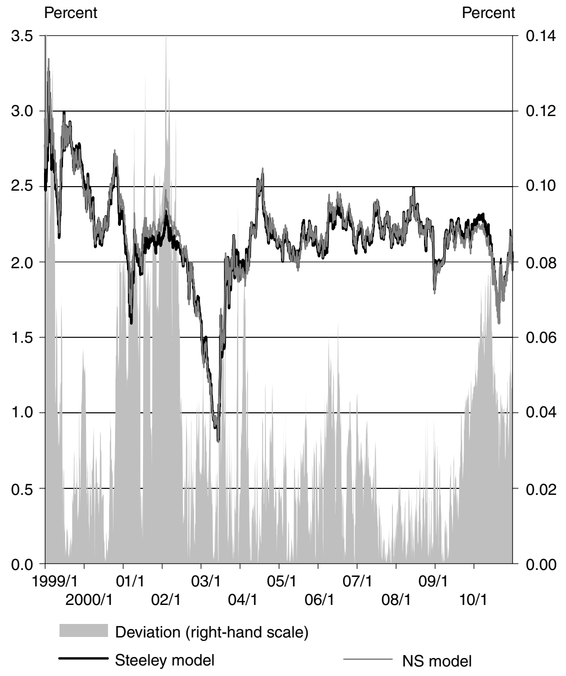
16. The six-month zero yields in this period using the Steeley model are a little higher than the zero yields derived from the short-term discount bonds issued by the Japanese government. This seems to reflect the fact that zero curves in this paper are estimated with only JGBs and our estimates of maturities of less than one year include some spreads compared to the zero yields of short-term discount bonds. In this paper, we do not use short-term discount bond prices for the zero curve estimation, because the market participants in the short-term discount bond market differ from those in the JGB market.

Figure 14 Comparison of Long-Term Interest Rate Estimates

[1] 10-Year Zero Yield



[2] 20-Year Zero Yield



B. Comparison of the Steeley Model with Smoothing Spline Methods That Can Estimate Smooth Instantaneous Forward Rates

An important point to note in using the piecewise polynomial method including the Steeley model is that the term structure of the instantaneous forward rate may have excessive unevenness. To overcome this, previous studies focusing on estimations of a smooth instantaneous forward rates term structure adopt what are known as smoothing spline methods (Fisher, Nychka, and Zervos [1995], Waggoner [1997], Jarrow, Ruppert, and Yu [2004], etc.). As seen below, smoothing spline methods estimate a smooth instantaneous forward rate term structure; nevertheless, the estimates tend to have lower fits to market prices compared with estimates obtained when smoothing is not used. As noted above, in this paper our purpose is to estimate a zero curve that accurately grasps the features of the JGB yield curve. Accordingly, we do not include smoothing spline methods in Section III because they do not meet the prerequisite of a good fit to the market prices.

We do compare zero curves and instantaneous forward rate term structures estimated using the Steeley model and the Jarrow, Ruppert, and Yu (2004) model (hereafter, the JRY model), one of the smoothing spline methods. We now outline the JRY model. In this approach, the instantaneous forward rate is modeled using a B-spline function based on equation (21). This method tries to estimate the zero curve with a good fit to market prices and a smoothed instantaneous forward rate curve by solving the following optimization problem with a penalty term:

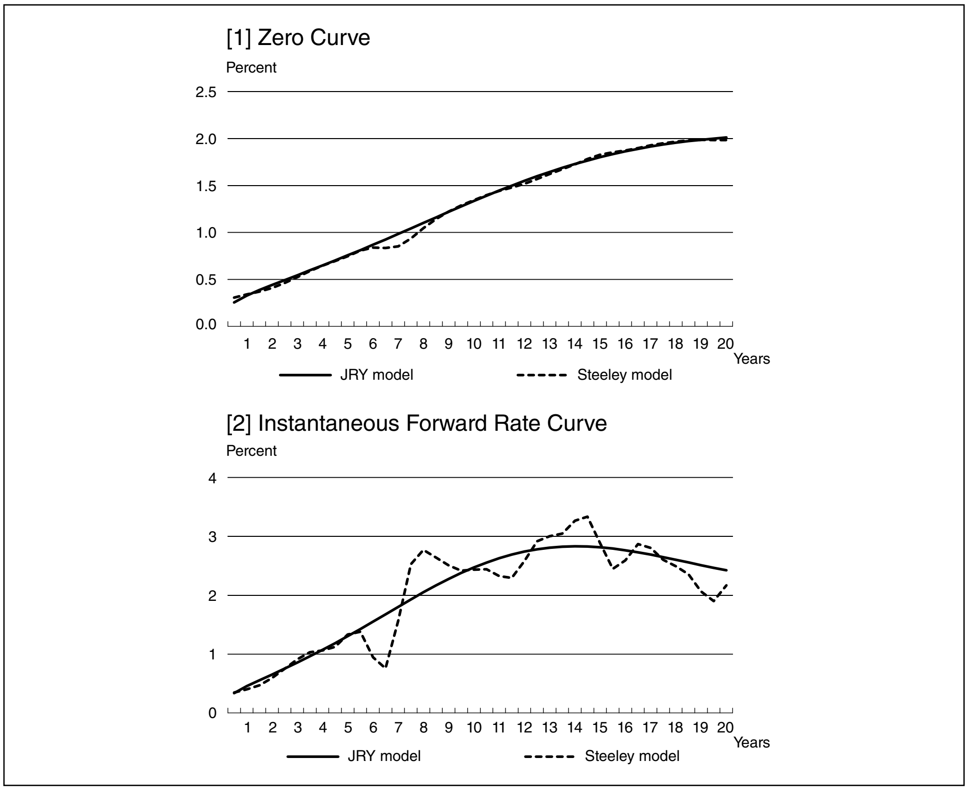
$$\min_{\boldsymbol{\alpha}} \left\{ \frac{1}{n_I}(\bar{\mathbf{P}} - \mathbf{Q}(\boldsymbol{\alpha}))^{\mathrm{T}}(\bar{\mathbf{P}} - \mathbf{Q}(\boldsymbol{\alpha})) + \lambda \int_0^{T_{cf}} (f(y)'')^2 dy \right\}. \tag{43}$$

The second term in equation (43) represents the penalty with respect to the curvature of the instantaneous forward rate. In other words, the objective function value rises as the unevenness of the instantaneous forward curve increases. Consequently, the optimal solution to equation (43) results in a zero curve estimate with a smooth instantaneous forward curve. The term λ is called the smoothing parameter. This parameter determines the unevenness of the instantaneous forward rate term structure. In Jarrow, Ruppert, and Yu (2004), this parameter is determined based on the information criteria proposed in Ruppert (1997).

Figure 15 shows the zero curves and the instantaneous forward rate term structures estimated with the Steeley model and the JRY model. For the zero curves, the Steeley model's curve shows a dip around the maturity of seven years, while this is not seen in the JRY model's curve. This implies that the Steeley model has a good fit to the market prices and grasps the characteristics of the JGB yield curve. On the other hand, the instantaneous forward rate curve based on the Steeley model shows great unevenness, while that of the JRY model is smooth.

In the financial business, problems in pricing derivatives underlying the forward interest rate occur when using an uneven instantaneous forward rate curve. The pricing of the derivatives is implemented by calculating the future forward curve based on some term structure model through the initial curve of the instantaneous forward rate. Hence,

Figure 15 Comparison of Steeley Model and JRY Model Estimation Results (February 17, 2009)



the price depends on the initial curve. Therefore, we end up deriving unreasonable prices for these derivatives if we use an uneven instantaneous forward curve as the initial curve of the term structure model.

To avoid this problem, we can price the derivatives underlying the forward interest rate with a smooth instantaneous forward curve estimated using a smoothing spline method. However, as pointed out above, zero curves estimated using smoothing spline methods have a worse fit to market prices. In this way, it is difficult to simultaneously achieve a good fit to JGB market prices and a smooth instantaneous forward rate term structure. Thus, we believe that the decision on whether to use the smoothing spline method or not should depend on how the estimated curve will be used.

V. Conclusion

In this paper, we set criteria to select an appropriate method from among estimation methods proposed in previous studies to grasp the characteristics of the JGB zero curve, and conducted a comparative analysis of the diverse methods according to the criteria. Specifically, we excluded estimation methods that did not meet the criteria that estimates should not fall below zero and take abnormal values. We then selected the most appropriate estimation method based on the criteria of a good fit to market prices and low unevenness in the zero curve. As a result, we found that the Steeley (1991) Method Modeling Discount Rates was the most appropriate in light of the criteria.

The zero yields estimated with the Steeley model did not drop below zero during the estimation period, and captured the characteristics frequently seen in the JGB yield curves including the flat shape near zero for short-term maturities and the inflection points in the curve for medium-term maturities. Zero curves estimated appropriately in this manner are considered to be important as a starting point for numerous analyses on JGB interest rates. Since 2010, the U.S. and European government bond yield curves have also been showing a flat shape near zero for short-term interest rates, coming to resemble the characteristics of the JGB yield curve during quantitative easing policy periods and in recent years. Thus, the Steeley model is likely to prove effective for zero curve estimations for the United States and European countries.

In conclusion, we would like to note two outstanding challenges for future research. First, the analysis in this paper is based on JGB data from the late 1990s through 2010. Hence, conducting a similar analysis using data from a different time or market might lead to different conclusions. Consequently, one challenge is to conduct a comparative analysis on zero curve estimation methods using more wide-ranging data. Second, the analysis in this paper uses the simple residual sum of squares of the market prices and theoretical prices as the objective function in estimating the parameters defining the zero curve. Depending on the market environment and the purpose of utilization of the zero curve, however, it may be more appropriate to conduct the estimation based on a different objective function. Therefore, another challenge for future research is to conduct the analysis using different objective functions.

APPENDIX 1: MARKET CONVENTIONS FOR CALCULATION OF JGB THEORETICAL PRICES

To precisely calculate the theoretical prices, it is necessary to understand JGB market conventions including the definition of the timing of when cash flow is paid, the method of calculating the number of days until cash flow is paid, and the method of calculating accrued interest. This appendix explains these items required for a precise calculation of theoretical prices.[17]

A. Definition of Terms

Here we present the terms required to explain JGB market conventions.

1. Interest payment dates

The dates on which coupons are paid on a bond are referred to as the interest payment dates. For fixed coupon-bearing bonds, these are normally set once every half-year, such as June 20 and December 20 each year. As explained in Appendix 1, Section B below, it is necessary to note that the interest payment dates sometimes differ from the dates on which coupons are actually paid. The interest payment dates are announced by the Ministry of Finance when JGB bond auctions are held. The interest payment

17. We referred to Ohta (2003) for an explanation of this appendix.

dates on the five-year, 10-year, 20-year, and 30-year bonds are all either March 20 and September 20 or June 20 and December 20 each year.[18] The interest payment dates on two-year bonds are the 20th of each month for bonds issued through September 2007, and the 15th of each month for bonds issued since October 2007.

2. Initial interest payment date

The first interest payment date after a given bond is issued, for the initial interest payment period, is referred to as the initial interest payment date. The initial interest payment date and the actual interest payment date sometimes differ.

3. Initial interest payment

The initial interest payment is the amount of cash flow obtained when the first interest payment occurs. Here we note that this is not the interest payment date.

4. Prior-term interest payment date

The prior-term interest payment date is defined as the date half a year prior to the initial interest payment date. For example, if the initial interest payment date is June 20, 2008, then the prior-term interest payment date is December 20, 2007.

5. Trade date

The trade date is the date on which the buyer and seller agree to the transaction.

6. Settlement date

The settlement date is the date three business days after the trade is executed, counting from the day after the trade date. The settlement date, not the trade date, is used as the standard for the calculation of occurrence of cash flow and accrued interest.

B. Market Conventions for Calculating the Number of Days until Cash Flow Is Paid

Here we explain the JGB market conventions for counting the number of days from the settlement date until the time when cash flow is paid. This is one of the factors used to calculate JGB theoretical prices.

First, we explain the case when the interest payment date is a business day. In this case, the interest payment date is the date when cash flow is actually paid. With 365 days in a year, the number of the days from the settlement date to the interest payment date is counted as the number of days until each interest payment date, including leap days for issues with less than a year remaining until the maturity date. For issues with a year or longer remaining until the maturity date, the count is the number of days from the settlement date to each interest payment date, deducting any leap days during this interval.

The one-end method is applied to the calculation of the number of days between the settlement date and interest payment dates. In this method, the settlement date is not counted in the calculation of the number of days, but each of the interest payment dates is counted. On the JGB market, the one-end method is applied to the day counting for all fixed coupon-bearing bonds.[19]

18. We note here that some of the 30-year bond issues have either February 20 and August 20 or May 20 and November 20 as the interest payment dates.
19. Prior to the introduction of the reopening system in March 2001 (see Appendix 1, Section C.1), the both-ends method was used for the calculation of accrued interest. This method counts both the settlement date and the coupon date in the day of counting.

In cases when the interest payment date is not a business day, the modified following business day convention is applied on the JGB market. Under this convention, when the interest payment date is not a business day and the subsequent business day is in the same month as the coupon date, the cash flow is paid on this date, and when the subsequent business day is in the month after the interest payment date, the cash flow is paid on the business day immediately preceding the interest payment date. As a specific example, many JGBs had interest payment dates on June 20, 2009. Because this date was a Saturday, the modified following business day convention was applied and the cash flow was paid on June 22, 2009.[20] In cases when the interest payment date is not a business day, the one-end method is also applied to the day counts between the settlement date and interest payment dates.

C. Response to Changes in Legal and Market Systems

Looking at the history of the JGB market, the times when cash flows were paid and the amount of cash flow changed together with developments in market and legal systems. As examples, we examine (1) the introduction of the immediate reopening rule, (2) the Happy Monday system, and (3) the settlement system.

1. Introduction of the immediate reopening rule (reopening system)

The immediate reopening rule (hereafter, reopening system) introduced in March 2001 is a system whereby upon issuance of a new bond with the same face value, coupon rate, and maturity date of an already-issued bond, the two bonds are treated as the same issue. Prior to the introduction of this system, the old and new bonds had to be treated as separate issues until the initial interest payment date even if the issuance conditions aside from the issuance date were identical, because the initial interest payments differed. The reopening system was introduced to reduce such complexities of managing issues.

The JGB initial interest payment calculation methods differ before and after the introduction of the reopening system. The accrued interest calculation methods also differ. The specific calculation methods are explained below.

2. The Happy Monday System

The Happy Monday System moved a number of public holidays to Mondays, to create three-day weekends. This system began under the Act on Partial Revision of the Act on National Holidays, which came into effect from January 2001. This moved the Coming of Age Day (January 15) and Health and Sports Day (October 10) holidays to the second Mondays of their respective months. Then, under the Act on Partial Revision of the Act on National Holidays and the Act on Social Welfare Service for the Elderly, which came into effect from January 2003, the Marine Day (July 20) and Respect for the Aged Day (September 15) holidays were moved to the third Mondays of their respective months.

The second revision had the greater effect from the perspective of calculating JGB theoretical prices. For example, July 20, which was a holiday prior to the revision, could now be a business day and the actual date when two-year JGB cash flow was paid changed because of the revision of the act. In addition, with the revision, September 20

20. The auction results for past interest-bearing JGBs are presented on the Ministry of Finance web page (http://www.mof.go.jp/english/jgbs/auction/past_auction_schedule/auct_resul/index.htm).

might now be Autumnal Equinox Day or Respect for the Aged Day. This could result in substitute holidays when Sundays and holidays overlap, potentially pushing the interest payment dates and the actual dates when cash flows are paid ahead or behind by a few days.[21]

In estimating the zero curve, it is necessary to specify exactly when such system changes were incorporated into market prices. Yet it is very difficult to accurately specify such information from the market prices.

For this reason, in our analysis in this paper, when calculating theoretical prices on trade dates from June 22, 2001 (when the Act on Partial Revision of the Act on National Holidays and the Act on Social Welfare Service for the Elderly was published in the government gazette), the dates when cash flows are paid after January 2003 (when that act came into effect) are assumed to be determined based on the revised holiday system.

3. Settlement system

In the JGB market, settlements among large brokerages switched to $T + 7$ (delivery seven business days later, counting from the day after the trade) from September 19, 1996 and then to $T + 3$ (delivery three business days later, counting from the day after the trade) from April 1, 1997. Because the calculation of theoretical JGB prices is based not on the trade date but on the settlement date, changes in the number of days between the trade date and the settlement date influence the theoretical price calculations. Because the market price data for our analysis in this paper run from January 1999 through December 2010, however, the above changes in the settlement system do not affect the JGB theoretical price calculations herein.

D. Initial Interest Payment and Accrued Interest Calculation Methods

As noted above, the initial interest payment and accrued interest calculation methods vary before and after the introduction of the reopening system in March 2001. Thus, we explain the calculation methods before and after the introduction of the reopening system.

1. Initial interest payment calculation method

Prior to the introduction of the reopening system, the initial interest payment for each issue was determined along with the issuance date as described below. As a result, issues with the same face value, coupon rate, and maturity but with different issuance dates had different initial interest payments.

The initial interest payment $g(c^{v_i}, N^{v_i}, T_1^{v_i})$ before the introduction of the reopening system is defined below.[22] When the issuance date of the bond preceded the prior-term interest payment date, the definition was as shown by the following equation with d as the number of days from the issuance date to the prior-term interest payment date using the both-ends method:

21. For example, in September 2009 the interest payment date was changed from the 21st to the 24th of the month because Respect for the Aged Day was moved to the third Monday. This was because the weekday between Respect for the Aged Day (September 21) and Autumnal Equinox Day (September 23) became a national holiday under the provisions of the Act on National Holidays.

22. In Appendix 1, Sections D.1 and 2, we assume that the prior-term interest payment date means the date after applying the modified following business day convention if it is not a business day.

$$g(c^{v_i}, N^{v_i}, T_1^{v_i}) = c^{v_i} N^{v_i} \left(\frac{1}{2} + \frac{d}{365}\right).$$

When the issuance date of the bond followed the prior-term interest payment date, the definition was as shown by the following equation with d as the number of days from the prior-term interest payment date to the date when the first interest payment occurred using the both-ends method:

$$g(c^{v_i}, N^{v_i}, T_1^{v_i}) = c^{v_i} N^{v_i} \frac{d}{365}.$$

For bonds issued since the introduction of the reopening system, the initial interest payment is defined as one-half the product of the coupon rate multiplied by the face value as shown by the following equation, regardless of the issuance date:

$$g(c^{v_i}, N^{v_i}, T_1^{v_i}) = \frac{c^{v_i} N^{v_i}}{2}.$$

2. Accrued interest calculation method

When a bond is traded between two interest payment dates, the seller has the right to receive cash flow corresponding to the number of days from the last day that cash flow was paid to the settlement date. However, all the cash flow for this period is paid to the bond holder when the subsequent cash flow is paid. Hence, the buyer pays this cash flow that the seller should receive by adding it to the market price when the bond is traded. This amount added is referred to as accrued interest. As mentioned above, in the JGB market the method of calculating accrued interest changed with the introduction of the reopening system. Accordingly, the following discussion is divided into the cases before and after the introduction of the reopening system in March 2001.

First, we present the accrued interest calculation method before the introduction of the reopening system. When the settlement date was prior to the date when the first interest payment occurred, the accrued interest A^{v_i} was calculated as follows with d' as the number of days from the issuance date to the settlement date according to the both-ends method, which includes leap days:

$$A^{v_i} = c^{v_i} N^{v_i} \times \frac{d'}{365}.$$

Before the introduction of the reopening system, when the settlement date was after the date when the first interest payment occurred, the accrued interest A^{v_i} was calculated as follows with d' as the number of days from the last date that cash flow was paid to the settlement date:

$$A^{v_i} = \begin{cases} c^{v_i} N^{v_i} \times \dfrac{d'}{365} & d' < 183 \\ \dfrac{c^{v_i} N^{v_i}}{2} & d' \geq 183 \end{cases}.$$

The method of calculating accrued interest after the introduction of the reopening system is divided into the two cases: (1) when the settlement date precedes the initial interest payment date; and (2) when the settlement date follows the initial interest payment date, as described below. In case (1), d' is the number of days from the settlement date to the prior interest payment date using the one-end method including leap days within the interval, and in case (2), d' is the number of days from the last day of occurrence of the cash flow to the settlement date using the one-end method including leap days within the interval. Then, the accrued interest is calculated using the following equation for both case (1) and case (2):

$$A^{v_i} = \begin{cases} c^{v_i} N^{v_i} \times \dfrac{d''}{365} & d'' < 183 \\ \dfrac{c^{v_i} N^{v_i}}{2} & d'' \geq 183 \end{cases}.$$

APPENDIX 2: ESTIMATION ALGORITHMS FOR THE STEELEY (1991) METHOD

In Appendix 1, we summarized the market conventions required to calculate JGB theoretical prices regarding the definition of when cash flow is paid, the method of calculating the number of days until cash flow is paid, and the determination of cash flow amounts (initial interest payments and accrued interest). When we estimate zero yields from JGB price data with the Steeley (1991) model selected in Section III, we have to make an estimate according to the calculations consistent with the market conventions explained in Appendix 1, as well as the theoretical aspects of the method presented in Section II. In Appendix 2, we now explain the algorithms for efficient conduct of estimations using the Steeley (1991) model while having the contents as explained in Appendix 1 accurately reflected in the JGB theoretical price calculations.

First we classify all JGBs traded at $t = 0$ into the following four groups:

(1) Issues with less than a year until maturity with interest payment dates on the 20th.
(2) Issues with less than a year until maturity with interest payment dates on the 15th.
(3) Issues with a year or more until maturity with interest payment dates on the 20th.
(4) Issues with a year or more until maturity with interest payment dates on the 15th.

Groups (2) and (4) cover all the two-year bonds issued after October 2007, as shown in Appendix 1, Section A. Groups (1) and (3) cover all other bond issues. The reason why we separate bonds that will mature within one year from all other bonds is that the adjustment of the leap year to calculate the number of days until the payment of cash flow differs according to this difference in the remaining maturities (see Appendix 1, Section B).

For simplicity, we show how to calculate the theoretical bond prices only for group (3) below. First, we set $t = 0$ as a certain day of month W in year V. The maturity date of the issue with the longest time to maturity among all issues in group (3) is assumed to be a certain day of month Y in year X. Then the number of months between these two dates, including both months to which both dates belong, is equal to $12(X - V) - W + Y + 1$. Hereafter this value is referred to as "n_{month}."

Next, we set the date on which issues in group (3) can pay cash flow as the 20th of each month from month W year V to month Y year X in cases when the 20th is a business day, and we set the date as the business day according to the modified following business day convention in cases when the 20th is not a business day. Then we calculate the number of years from $t = 0$ to each of these dates based on the method described in Appendix 1, Section B. We set these as $T_1 < T_2 < \cdots < T_{n_{month}}$.

Additionally, if issues in group (3) at $t = 0$ are expressed as $v_i (i = 1, \ldots n)$, then the amount of cash flow for issue v_i on the 20th of each month (or on the date determined by the modified following business day convention in cases when the 20th is not a business day) from month W year V to month Y year X is determined. We denote these cash flows for issue v_i as $\tilde{\mathbf{c}}^{v_i} := (\tilde{c}_1^{v_i}, \ldots, \tilde{c}_{n_{month}}^{v_i})^{\mathrm{T}}$. All JGB issues generate interest payments every half-year; hence, it is important to note that some of elements of $\tilde{\mathbf{c}}^{v_i}$ take zero.

To calculate all discount functions until dates when cash flows are paid, we define matrix $\mathbf{B}$ with an element taking the value of a B-spline function for all points of time below. We define the (k, j)-th element of matrix $\mathbf{B}$ as $B(j - 4, T_k)$ where $B(k, x)$ is calculated from equation (16) introduced in Section II.D.1.a.(2). In other words, we define matrix $\mathbf{B}$ as shown in equation (A.1).

$$\mathbf{B} = \begin{pmatrix} B(-3, T_1) & \cdots & B(n_{knot} - 1, T_1) \\ \vdots & \ddots & \vdots \\ B(-3, T_{n_{month}}) & \cdots & B(n_{knot} - 1, T_{n_{month}}) \end{pmatrix}. \tag{A.1}$$

Here, we find that the discount rate $Z(T_j)$ for T_j becomes the j-th element of $\mathbf{B}\boldsymbol{\alpha}$ from equation (15) if we set the parameter of the Steeley (1991) method as $\boldsymbol{\alpha} = (\alpha_{-3}, \ldots, \alpha_{n_{knot}-1})^{\mathrm{T}}$.

Thus, the theoretical price of issue v_i, $Q^{v_i}(\boldsymbol{\alpha})$ is calculated as shown in equation (A.2).

$$Q^{v_i}(\boldsymbol{\alpha}) = (\tilde{\mathbf{c}}^{v_i})^{\mathrm{T}} \mathbf{B}\boldsymbol{\alpha}. \tag{A.2}$$

The efficiency of this algorithm is that when we calculate the discount function, it is not necessary to calculate the number of days until each cash flow is paid for each issue and a B-spline function for each day. In other words, within the same group of issues—the group (3) issues—we have only to calculate the **B** matrix just once following equation (A.1).

In the above, we present the calculation algorithm for the theoretical prices of group (3) issues. As in the above case, applying the same algorithm for the issues in the other groups—groups (1), (2), and (4)—we can calculate the theoretical prices following equation (A.2) by calculating the **B** matrix just once for each group from equation (A.1). Thus, this algorithm enables us to estimate the zero curve efficiently by calculating the **B** matrix for each of the four groups rather than calculating the spline functions for the cash flow paid at each point in time for each issue.

References

Bank for International Settlements, "Zero-Coupon Yield Curves: Technical Documentation," BIS Paper No. 25, Bank for International Settlements, 2005.

Carleton, W. T., and I. A. Cooper, "Estimation and Uses of the Term Structure of Interest Rates," *Journal of Finance*, 31 (4), 1976, pp. 1067–1083.

Fisher, M., D. Nychka, and D. Zervos, "Fitting the Term Structure of Interest Rates with Smoothing Splines," working paper, Board of Governors of the Federal Reserve System, 1995.

Gürkaynak, R. S., B. Sack, and J. H. Wright, "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, 54 (8), 2007, pp. 2291–2304.

Houglet, M. X., "Estimating the Term Structure of Interest Rates for Non-Homogeneous Bonds," Ph.D. Dissertation, University of California Berkeley, 1980.

Inui, K., and Y. Muromachi, *Kin'yu Moderu ni okeru Suitei to Saitekika* (Estimation and Optimization in Financial Modeling), Asakura Shoten, 2000 (in Japanese).

Ioannides, M., "A Comparison of Yield Curve Estimation Techniques Using UK Data," *Journal of Banking and Finance*, 27 (1), 2003, pp. 1–26.

Jarrow, R., D. Ruppert, and Y. Yu, "Estimating the Interest Rate Term Structure of Corporate Debt with a Semiparametric Penalized Spline Model," *Journal of the American Statistical Association*, 99 (465), 2004, pp. 57–66.

Kalev, P. S., "Estimating and Interpreting Zero Coupon and Forward Rates: Australia, 1992–2001," working paper, Monach University, 2004.

Kawasaki, Y., and T. Ando, "Estimating Term Structure Using Non-Linear Splines: A Penalized Likelihood Approach," in A. Zerger and R. M. Argent, eds. *Advances in Applications for Management and Decision Making*, 2005, pp. 864–870.

Komine, M., M. Yamagishi, K. Matsumoto, T. Futatsugi, J. Tsukasa, T. Nagao, K. Sunagawa, and N. Sano, "Wagakuni Saiken Shijo no Koyu no Gensho to Kikan Kozo Bunseki (Analyses of Japan's Bond Market and Its Term Structure)," *Finansharu Rebyu* (Financial Review), 14, Ministry of Finance, 1989 (in Japanese).

McCulloch, J. H., "Measuring the Term Structure of Interest Rates," *Journal of Business*, 44 (1), 1971, pp. 19–31.

———, "The Tax-Adjusted Yield Curve," *Journal of Finance*, 30 (3), 1975, pp. 811–830.

———, and L. A. Kochin, "The Inflation Premium Implicit in the US Real and Nominal Term Structures of Interest Rates," working paper, Ohio State University, 2000.

Nelson, C. R., and A. F. Siegel, "Parsimonious Modeling of Yield Curves," *Journal of Business*, 60 (4), 1987, pp. 473–489.

Oda, N., "A Note on the Estimation of Japanese Government Bond Yield Curves," IMES Discussion Paper No. 96-E-27, Institute for Monetary and Economic Studies, Bank of Japan, 1996.

Ohta, T., *Shin Saiken Unyo to Toshi Senryaku* (Bond Investment and Investment Strategy), Kin'yu Zaisei Jijo Kenkyukai (Kinzai Institute for Financial Affairs), 2003 (in Japanese).

Ruppert, D., "Empirical-Bias Bandwidths for Local Polynomial Nonparametric Regression and Density Estimation," *Journal of the American Statistical Association*, 92 (439), 1997, pp. 1049–1062.

Schaefer, S. M., "Measuring a Tax-Specific Term Structure of Interest Rates in the Market for British Government Securities," *Economic Journal*, 91 (362), 1981, pp. 415–438.

Steeley, J. M., "Estimating the Gilt-Edged Term Structure: Basis Splines and Confidence Intervals," *Journal of Business Finance and Accounting*, 18 (4), 1991, pp. 513–529.

Svensson, L. E. O., "Estimating Forward Interest Rates with the Extended Nelson and Siegel Method," *Sveriges Riksbank Quarterly Review*, 3 (1), 1995, pp. 13–26.

Tanggaard, C., "Nonparametric Smoothing of Yield Curves," *Review of Quantitative Finance and Accounting*, 9 (3), 1997, pp. 251–267.

Vasicek, O. A., and H. G. Fong, "Term Structure Modeling Using Exponential Splines," *Journal of Finance*, 37 (2), 1982, pp. 339–348.

Waggoner, D. F., "Spline Methods for Extracting Interest Rate Curves from Coupon Bond Prices," Federal Reserve Bank of Atlanta Working Paper No. 97-10, Federal Reserve Bank of Atlanta, 1997.